

# Optimal Adaptive Random Multiaccess in Energy Harvesting Wireless Sensor Networks

Nicolò Michelusi, *Member, IEEE*, and Michele Zorzi, *Fellow, IEEE*

**Abstract**—Wireless sensors can integrate rechargeable batteries and energy-harvesting (EH) devices to enable long-term, autonomous operation, thus requiring intelligent energy management to limit the adverse impact of energy outages. This work considers a network of EH wireless sensors, which report packets with a random utility value to a fusion center (FC) over a shared wireless channel. Decentralized access schemes are designed, where each node performs a *local* decision to transmit/discard a packet, based on an estimate of the packet's utility, its own energy level, and the scenario state of the EH process, with the objective to maximize the average long-term aggregate utility of the packets received at the FC. Due to the non-convex structure of the problem, an approximate optimization is developed by resorting to a mathematical artifice based on a game theoretic formulation of the multiaccess scheme, where the nodes do not behave strategically, but rather attempt to maximize a *common* network utility with respect to their own policy. The *symmetric Nash equilibrium* (SNE) is characterized, where all nodes employ the same policy; its uniqueness is proved, and it is shown to be a local maximum of the original problem. An algorithm to compute the SNE is presented, and a heuristic scheme is proposed, which is optimal for large battery capacity. It is shown numerically that the SNE typically achieves near-optimal performance, within 3% of the optimal policy, at a fraction of the complexity, and two operational regimes of EH-networks are identified and analyzed: an *energy-limited scenario*, where energy is scarce and the channel is under-utilized, and a *network-limited scenario*, where energy is abundant and the shared wireless channel represents the bottleneck of the system.

## I. INTRODUCTION

In recent years, energy harvesting (EH) has received great interest in the research community, spanning different disciplines related to electronics, energy storage, measurement and modeling of ambient energy sources [2], and energy management [3], [4]. EH devices can operate autonomously over long periods of time, as they are capable of collecting energy from the surrounding environment, and of employing it to sustain tasks such as data acquisition, processing and transmission. Due to its demonstrated advantages of long-term, autonomous operation, at least within the limits of the physical degradation of the rechargeable batteries [5], the EH technology exploiting, *e.g.*, solar, motion, heat, aeolian harvesting [6], or indoor light [7] is increasingly being considered in the design of wireless sensor networks (WSNs), where battery replacement is difficult or cost-prohibitive, *e.g.*, see [8].

The rechargeable battery of an EH sensor (EHS) can be modeled as an *energy buffer*, where energy is stored according to a given statistical process, and from where energy is drawn to feed sensor microprocessors and transceiver equipments, whenever needed. However, due to the random and uncertain energy supply, *battery management* algorithms are needed to manage the harvested energy, in order to minimize the deleterious impact of energy outage, with the goal of optimizing the long-term performance of the overall system in regard to sensing and data-communication tasks [3], [9]–[12]. In the literature, different battery management algorithms have been presented, exploiting knowledge about, *e.g.*, the amount of energy stored in the battery, which may be perfectly [13], [14] or imperfectly [15] known to the controller, the importance of the data packets to be transmitted [16], the state of the EH source [16], the health state of the battery [5], or the channel state [17].

In this paper, extending our previous work [1], we consider an EH-WSN where multiple EHSs randomly access a wireless channel to transmit data packets with random utility value to a FC. The goal is to design strategies so as to maximize the long-term average utility of the data reported to the FC. One possibility to achieve this goal is to have the FC act as a centralized controller, which schedules each EHS based on knowledge of the energy levels and packet utilities of all EHSs. This problem has been considered for the case of two EHSs in [18]. While this centralized scheme is expected to yield the best performance, since collisions between concurrent transmissions can be avoided via proper scheduling at the FC, it can only be used in very small WSNs, such as the one considered in [18]. In fact, the EHSs need to report their energy level and packet's utility to the FC in each slot, thus incurring a very large communication overhead in large WSNs, which is not practical given the limited resources (shared channel and ambient energy) of EH-WSNs. Therefore, unlike [18], we consider a fully decentralized scheme with an arbitrary number of EHSs, where the EHSs operate without the aid of a central controller. In particular, their decision as to whether to transmit the data packet or to remain idle and discard it is based only on *local* information about their own energy level and an estimate of the utility of the current data packet, rather than on *global* network state information. Compared to the centralized scheme, the decentralized one achieves poorer performance, due to the unavoidable collisions between concurrent EHSs, but is scalable to large EH-WSNs, since it does not incur communication overhead between the EHSs and the FC, but only needs minimal coordination by the FC, which, from time to time (*i.e.*, whenever environmental conditions change), needs to broadcast information about the

The research of Nicolò Michelusi has been funded in part by the following grants: AFOSR FA9550-12-1-0215, ONR N00014-09-1-0700.

N. Michelusi is with the Dept. of Electrical Engineering, University of Southern California. email: michelus@usc.edu.

M. Zorzi is with the Dept. of Information Engineering, University of Padova, Padova, Italy. email: zorzi@dei.unipd.it.

Part of this work was presented at IEEE ICC 2013 [1].

policy to be employed by the EHSs to perform their access decisions. Assuming that data transmission incurs an energy cost and that simultaneous transmissions from multiple EHSs cause collisions and packet losses, we study the problem of designing optimal decentralized random access policies, which manage the energy available in the battery so as to maximize the network utility, defined as the average long-term aggregate network utility of the data packets successfully reported to the FC. In order to tackle the non-convex nature of the optimization problem, we resort to a mathematical artifice based on a game theoretic formulation of the multiaccess problem, where each EHS tries to individually maximize the network utility. Owing to the symmetry of the network, we characterize the symmetric Nash equilibrium of this game, such that all EHSs employ the same scheme to access the channel, and prove its existence and uniqueness. We also show that it is a local optimum of the original optimization problem, and derive an algorithm to compute it, based on the bisection method and policy iteration. Moreover, we design a heuristic scheme, which is proved to be asymptotically optimal for large battery capacity, based on which we identify two regimes of operation of EH-WSNs: an *energy-limited scenario*, where the bottleneck of the system is due to the scarce energy supply, resulting in under-utilization of the shared wireless channel; and a *network-limited scenario*, where energy is abundant and the network size and the shared wireless channel are the main bottlenecks of the system.

The model considered in the present work is a generalization of that of [19], which addresses the design of optimal energy management policies for a single EHS. Herein, we extend [19] and model the interaction among multiple EHSs, which randomly access the channel. The problem of maximizing the average long-term importance (similar to the concept of utility used in this work) of the data reported by a replenishable sensor is formulated in [20] for a continuous-time model with Poisson EH and data processes, whereas [21] investigates the relaying of packets of different priorities in a network of energy-limited sensors (no EH).

Despite the intense research efforts in the design of optimal energy management policies for a single EH device, *e.g.*, see [14], [22], [23], the problem of analyzing and modeling the interaction among multiple EH nodes in a network has received limited attention so far. In [24], [25], the design of medium access control (MAC) protocols for EH-WSNs is considered, focusing on TDMA, (dynamic) framed Aloha, CSMA and polling techniques. In [26], efficient energy management policies that stabilize the data queues and maximize throughput or minimize delay are considered, as well as efficient policies for contention free and contention based MAC schemes, assuming that the data buffer and the battery storage capacity are infinite. In contrast, we develop decentralized random multiaccess schemes under the practical assumption of *finite* battery storage capacity. In [27], a MAC protocol is designed using a probabilistic polling mechanism that adapts to changing EH rates or node densities. In [28], an on-demand MAC protocol is developed, able to support individual duty cycles for nodes with different energy profiles, so as to achieve energy neutral operation. In [10], a dynamic sensor activation

framework to optimize the event detection performance of the network is developed. In [29], the problem of maximizing the lifetime of a network with finite energy capacity is considered, and a heuristic scheme is constructed that achieves close-to-maximum lifetime, for battery-powered nodes without EH capability. In [30], the problem of joint energy allocation and routing in multihop EH networks in order to maximize the total system utility is considered, and a low-complexity online scheme is developed that achieves asymptotic optimality, as the battery storage capacity becomes asymptotically large, without prior knowledge of the EH profile. In [13], [31], [32], tools from Lyapunov optimization and weight perturbation are employed to design near-optimal schemes to maximize the network utility in multihop EH networks. In particular, [13] considers the problem of cross-layer resource allocation for wireless multihop networks operating with rechargeable batteries under general arrival, channel state and recharge processes, and presents an online and adaptive policy for the stabilization and optimal control of EH networks. In [31], the problem of dynamic joint compression and transmission in EH multihop networks with correlated sources is considered. In [32], the problem of joint energy allocation and data routing is considered, and near-optimal online power allocation schemes are developed. However, [32] does not consider wireless interference explicitly, and is thus applicable to settings where adjacent nodes operate on orthogonal channels. In contrast, we employ a collision channel model, where the simultaneous transmission of multiple nodes results in packet losses.

The rest of the paper is organized as follows. In Sec. II, we present the system model. Sec. III defines the control policies and states the optimization problem, which is further developed in Sec. IV. In Sec. V, we present and analyze the performance of a heuristic policy. In Sec. VI, we present some numerical results. Finally, in Sec. VII we conclude the paper. The proofs of the propositions and theorems are provided in the Appendix.

## II. SYSTEM MODEL

We consider a network of  $U$  EHSs, which communicate concurrently via a shared wireless link to a FC, as depicted in Fig. 1. Each EHS harvests energy from the environment, and stores it in a rechargeable battery to power the sensing apparatus and the RF circuitry. A processing unit, *e.g.*, a micro-controller, manages the energy consumption of the EHS. Time is slotted, where slot  $k$  is the time interval  $[kT, kT + T)$ ,  $k \in \mathbb{Z}^+$  and  $T$  is the time-slot duration.

### A. Data acquisition model

At each time instant  $k$ , each EHS (say,  $u$ ) has a new data packet to send to the FC. Each data packet has utility  $V_{u,k} > 0$ . We model  $V_{u,k}$  as a continuous random variable with probability density function (pdf)  $f_V(v)$ ,  $v \geq 0$  with support  $(0, \infty)$ , and assume that the components of  $\{\mathbf{V}_k\}$ , where  $\mathbf{V}_k = (V_{1,k}, V_{2,k}, \dots, V_{U,k}) \in [\mathbb{R}^+]^U$  is the *utility* vector, are i.i.d. over time and across the EHSs. EHS  $u$  may not know exactly the utility  $V_{u,k}$  of the current data packet, but is only provided with a noisy observation  $Y_{u,k}$

Table I  
MAIN SYSTEM PARAMETERS

$e_{\max}$	Battery storage capacity	$E_{u,k} \in \mathcal{E}$	Energy level of EHS $u$ , time $k$
$B_{u,k} \sim \mathcal{B}(\beta(S_k))$	EH arrival at EHS $u$ , time $k$	$Q_{u,k} \in \{0, 1\}$	Transmit action at EHS $u$ , time $k$
$S_k \sim p_S(\cdot S_{k-1})$	Common scenario process	$\beta(s)$	Average EH rate in scenario $s$
$V_{u,k} \sim f_V(\cdot)$	Packet's utility at EHS $u$ , time $k$	$Y_{u,k} \sim f_{Y V}(\cdot V_{u,k})$	Packet's utility observation at EHS $u$ , time $k$
$\gamma_{u,k}$	Channel SNR for EHS $u$ , time $k$	$V(Y_{u,k})$	Expected utility for EHS $u$ , if it transmits at time $k$
$\eta_u(E_{u,k}, S_k)$	Transmission probability for EHS $u$	$g(\eta_u(E_{u,k}, S_k))$	Expected instantaneous utility under threshold policies
$R_\eta$	Network utility	$P(\eta_u s)$	Average power consumption in scenario $s$ for EHS $u$
$y_{\text{th},u}(e, s)$	Censoring threshold on $Y_{u,k}$	$\rho(y)$	Channel outage probability, given $Y_{u,k} = y$
$G(\eta_u s)$	Average utility in scenario $s$ for EHS $u$ under no collisions		

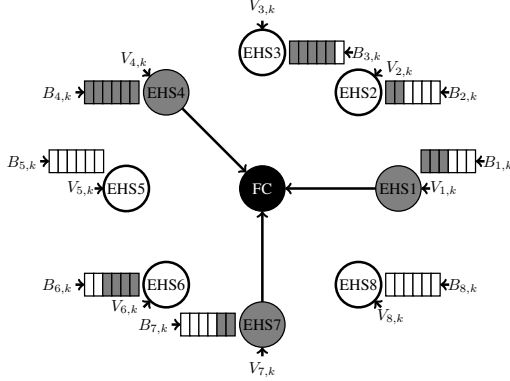


Figure 1. Energy Harvesting Wireless Sensor Network (EH-WSN)

of it, termed the *utility observation*, distributed according to the conditional probability density function  $f_{Y|V}(y|v)$  with support  $(0, \infty)$ ,  $\forall v$ . One example is when the utility  $V_{u,k}$  represents the achievable information rate under the current channel fading state. In this case,  $V_{u,k}$  is known only to some extent, via proper channel estimation. We denote the joint probability density function of  $(V_{u,k}, Y_{u,k})$  as  $f_{V,Y}(v, y)$ , the conditional probability density function of  $V_{u,k}$  given  $Y_{u,k}$  as  $f_{V|Y}(v|y)$ , and the marginal probability density function of  $Y_{u,k}$  as  $f_Y(y)$ .

### B. Channel model

We employ a collision channel model with fading,<sup>1</sup> i.e., letting  $\gamma_{u,k}$  be the SNR between EHS  $u$  and the FC in slot  $k$ , the transmission of EHS  $u$  in slot  $k$  is successful if and only if all the other EHSs remain idle and  $\gamma_{u,k} > \gamma_{\text{th}}(Y_{u,k})$ , where  $\gamma_{\text{th}}(Y_{u,k})$  is an SNR threshold, possibly dependent on the utility observation  $Y_{u,k}$ , which guarantees correct decoding at the receiver. The dependence on  $Y_{u,k}$  models the fact that the transmission rate may be adapted over time, based on  $Y_{u,k}$ . We model  $\{\gamma_{u,k}\}$  as a continuous random variable in  $(0, \infty)$ , i.i.d. over time and across EHSs, and possibly correlated with  $(V_{u,k}, Y_{u,k})$ . For instance, if  $V_{u,k}$  represents the achievable information rate under the current channel fading state, then  $V_{u,k} = \log_2(1 + \gamma_{u,k})$  [34], and  $Y_{u,k}$  is a noisy estimate of the achievable rate, obtained from a noisy estimate of the SNR  $\gamma_{u,k}$ . We denote the joint probability distribution of  $(V_{u,k}, Y_{u,k}, \gamma_{u,k})$  as  $f_{V,Y,\Gamma}(v, y, \gamma)$ , and we let  $\rho(y) \triangleq \mathbb{P}(\gamma_{u,k} \leq \gamma_{\text{th}}(y)|Y_{u,k} = y)$  be the channel outage probability, which depends on the specific fading model and

rate adaptation employed. Thus, letting  $O_{u,k} \in \{0, 1\}$  be the transmission outcome for EHS  $u$ , with  $O_{u,k} = 1$  if the transmission is successful and  $O_{u,k} = 0$  otherwise, and letting  $Q_{n,k} \in \{0, 1\}$  be the transmission decision of EHS  $n$ , with  $Q_{n,k} = 1$  if it transmits and  $Q_{n,k} = 0$  if it remains idle, the success probability under a specific transmission pattern  $\mathbf{Q}_k = (Q_{1,k}, Q_{2,k}, \dots, Q_{U,k}) \in \{0, 1\}^U$  and  $Y_{u,k} = y$  is given by

$$\mathbb{P}(O_{u,k}=1|\mathbf{Q}_k=\mathbf{q}, Y_{u,k}=y) = (1-\rho(y))q_u \prod_{i \neq u} (1-q_i). \quad (1)$$

The data packet is discarded if a collision occurs or the EHS decides to remain idle, and no feedback on the transmission outcome is provided by the FC to the EHSs. This assumption implicitly models a stringent delay requirement. Moreover, we do not deal explicitly with network management issues, e.g., node synchronization and self organization, nodes joining and leaving the network, fault tolerance, which can be addressed using techniques available in the literature.

Given the utility observation  $Y_{u,k} = y$ , EHS  $u$  can estimate the utility of the current data packet via conditional expectation, i.e.,

$$\bar{V}(y) \triangleq \mathbb{E}[\chi(\gamma_{u,k} > \gamma_{\text{th}}(Y_{u,k}))V_{u,k}|Y_{u,k} = y]. \quad (2)$$

Note that  $\bar{V}(y)$  represents the expected utility for EHS  $u$ , if the packet is transmitted, and takes into account possible channel outage events due to channel fading. In the special case where  $\gamma_{u,k}$  and  $V_{u,k}$  are mutually independent, given  $Y_{u,k}$ , we obtain  $\bar{V}(y) = (1 - \rho(y))\mathbb{E}[V_{u,k}|Y_{u,k} = y]$ . We make the following assumptions on  $\bar{V}(y)$ .

**Assumption 1**  $\bar{V}(y)$  is a strictly increasing, continuous function of  $y \in (0, \infty)$ , with  $\lim_{y \rightarrow \infty} \bar{V}(y) = \infty$  and  $\bar{V}(0) = 0$ .

**Remark 1** The joint distribution of  $(V_{u,k}, Y_{u,k}, \gamma_{u,k})$  is application/deployment dependent. Moreover, the assumption that  $V_{u,k}$ ,  $Y_{u,k}$  and  $\gamma_{u,k}$  have continuous distributions with support  $(0, \infty)$  approximates the scenario where they are mixed or discrete random variables as well. This is the case, for instance, if the data traffic is bursty, i.e., a data packet is received with probability  $\lambda$ , otherwise no data packet is received. In this case, the utility is binary and takes value  $V_{u,k} = 1$  if a packet is received and  $V_{u,k} = 0$  otherwise, whereas the utility observation is given by  $Y_{u,k} = V_{u,k}$ . The approximation for  $V_{u,k}$  is obtained as follows: let  $V_{u,k}$  be a mixed random variable, i.e., with probability  $x$ , it takes a value from the discrete set  $\mathcal{V}_D$  with probability mass function  $f_V^{(D)}(v)$ ,  $v \in \mathcal{V}_D$ , and, with probability  $1 - x$ , it takes a value

<sup>1</sup>The extension to the case of multiple orthogonal channels can be done by exploiting the large network approximation, similar to [33, Corollary 1]

from the continuous set  $\mathcal{V}_C$  with probability density function  $f_V^{(C)}(v)$  (in the discrete case,  $x = 1$ ). Then,  $V_{u,k}$  can be approximated as a continuous random variable with support  $(0, \infty)$  and probability density function

$$\tilde{f}_V^\sigma(v) = (1-x)f_V^{(C)}(v) + x \sum_{\nu_D \in \mathcal{V}_D} f_V^{(D)}(\nu_D) \frac{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(v-\nu_D)^2}{2\sigma^2}}}{1-\Phi(\nu_D/\sigma)}, v \geq 0,$$

where  $\Phi(x) = \mathbb{P}(N < x)$  is the cumulative distribution function of a Gaussian normal random variable  $N \sim \mathcal{N}(0, 1)$ . This approximation has support  $(0, \infty)$ , and approaches the true mixed distribution as  $\sigma \rightarrow 0$ . The same consideration holds for the case where the observation  $Y_{u,k}$  and the fading state  $\gamma_{u,k}$  have a discrete or mixed distribution. Therefore, without loss of generality, we restrict our analysis to the case where  $V_{u,k}$  and  $Y_{u,k}$  are continuous random variables with support  $(0, \infty)$ .

### C. Battery dynamics

The battery of each EHS is modeled by a buffer. As in previous works [19], [22], [23], we assume that each position in the buffer can hold one energy quantum and that the transmission of one data packet requires the expenditure of one energy quantum. In this work, we assume that no power adaptation is employed at the EHSs (however, the transmission rate may be adapted to the utility observation  $Y_{u,k}$ , see Sec. II-B). Moreover, we only consider the energy expenditure associated with RF transmission, and neglect the one due to battery leakage effects and storage inefficiencies, or needed to turn on-off the circuitry, etc. These aspects can be handled by extending the EH/consumption model, as discussed in [5], but their inclusion does not bring additional insights on the multiaccess problem considered in this paper, and therefore is not pursued here. The maximum number of quanta that can be stored in each EHS, *i.e.*, the battery capacity, is  $e_{\max}$  and the set of possible energy levels is denoted by  $\mathcal{E} = \{0, 1, \dots, e_{\max}\}$ . The amount of energy stored in the battery of EHS  $u$  at time  $k$  is denoted as  $E_{u,k}$ , and evolves as

$$E_{u,k+1} = \min \{E_{u,k} - Q_{u,k} + B_{u,k}, e_{\max}\}, \quad (3)$$

where  $\{B_{u,k}\}$  is the *energy arrival process* and  $\{Q_{u,k}\}$  is the *transmit action process* at EHS  $u$ .  $Q_{u,k} = 1$  if the current data packet is transmitted by EHS  $u$ , which results in the expenditure of one energy quantum, and  $Q_{u,k} = 0$  otherwise, as described in Sec. III.  $B_{u,k}$  models the randomness in the energy harvested in slot  $k$ , and is described in Sec. II-D. Moreover the energy harvested in time-slot  $k$  can be used only in a later time-slot, so that, if the battery is depleted, *i.e.*,  $E_{u,k} = 0$ , then  $Q_{u,k} = 0$ . The state of the system at time  $k$  is given by  $(\mathbf{E}_k, \mathbf{V}_k)$ , where  $\mathbf{E}_k = (E_{1,k}, E_{2,k}, \dots, E_{U,k}) \in \mathcal{E}^U$  is the joint state of the energy levels in all EHSs.

### D. Energy harvesting process

We model the energy arrival process  $\{B_{u,k}\}$  as a homogeneous hidden Markov process, taking values in the set  $\mathcal{B}$ . We define an underlying *common scenario process*  $\{S_k\}$ ,

taking values in the finite set  $\mathcal{S}$ , which evolves according to a stationary irreducible Markov chain with transition probability  $p_S(s_{k+1}|s_k) \triangleq \mathbb{P}(S_{k+1}=s_{k+1}|S_k=s_k)$ . Given the common scenario  $S_k = s$ , the energy harvest at EHS  $u$  is modeled as  $B_{u,k} \sim \mathcal{B}(\beta(s))$ , *i.e.*, in each time-slot, either one energy quantum is harvested with probability  $\beta(s)$ , or no energy is harvested at all. The common scenario process models different environmental conditions affecting the EH mechanism, for instance, for a solar source,  $S_k \in \{\text{direct sunlight, cloudy, night}\}$ . In this case, we thus have that  $\beta(\text{direct sunlight}) > \beta(\text{cloudy}) > \beta(\text{night})$ . We assume that the components of  $\mathbf{B}_k = (B_{1,k}, B_{2,k}, \dots, B_{U,k})$  are i.i.d. over time and across the EHSs, given the common scenario  $S_k = s$ . We define  $\pi_S(s)$ ,  $s \in \mathcal{S}$ , as the steady state distribution of the common scenario process, and we refer to  $\bar{\beta} = \mathbb{E}[B_{u,k}] = \sum_{s \in \mathcal{S}} \pi_S(s) \beta(s)$  as the *average EH rate*, and to  $\beta(s)$  as the *average EH rate in scenario*  $S_k = s$ . This model is a special instance of the *generalized Markov model* presented in [35] for the case of one EHS. Therein, the scenario process is modeled as a first-order Markov chain, whereas  $B_{u,k}$  statistically depends on  $(B_{u,k-L}, B_{u,k-L+1}, \dots, B_{u,k-1})$  and on  $S_k$ , for some order  $L \geq 0$ . In particular, in [35] it is shown that  $L = 0$  ( $\{B_{u,k}\}$  is i.i.d., given  $S_k$ ) models well a piezoelectric energy source, whereas  $L = 1$  ( $\{B_{u,k}\}$  is a Markov chain, given  $S_k$ ) models well a solar energy source. In this paper, the model  $L = 0$  is considered. The analysis can be extended to the case  $L = 1$  by including  $B_{k-1}$  in the state space, similarly to [5], but this extension is beyond the scope of this paper. Typically, the scenario process  $\{S_k\}$  is slowly varying over time, hence it can be estimated accurately at each EHS from the locally harvested sequence  $\{B_{u,k}\}$ . In this paper, we assume that perfect knowledge of  $S_k$  is available at each EHS at time instant  $k$ . Note that the common scenario process  $\{S_k\}$  introduces *spatial and temporal correlation* in the EH process, *i.e.*, the local EH processes are spatially coupled across the network since  $S_k$  is common to all EHSs and time-correlated since  $S_k$  is a Markov chain.

## III. POLICY DEFINITION AND OPTIMIZATION PROBLEM

Each EHS is assumed to have only *local* knowledge about the state of the system. Namely, at time  $k$ , EHS  $u$  only knows its own energy level  $E_{u,k}$ , the utility observation  $Y_{u,k}$ , and the common scenario state  $S_k$ , but does not know the energy levels and utilities of the other EHSs in the network, nor the exact utility of its own packet  $V_{u,k}$  or channel SNR  $\gamma_{u,k}$ , since these cannot be directly measured by EHS  $u$ . Moreover, since no feedback is provided by the FC, EHS  $u$  has no information on the collision history. Therefore, the transmission decision of EHS  $u$  at time  $k$  is based solely on  $(E_{u,k}, Y_{u,k}, S_k)$ .

Given  $(E_{u,k}, Y_{u,k}, S_k)$ , EHS  $u$  transmits the current data packet ( $Q_{u,k} = 1$ ) with probability  $\mu_u(E_{u,k}, Y_{u,k}, S_k)$ , and discards it otherwise ( $Q_{u,k} = 0$ ), for some transmission policy  $\mu_u$ , which is the objective of our design. Note that, due to the complexity of implementing a non-stationary policy  $\mu_{u,k}$ , we only consider stationary policies independent of time  $k$ . Let  $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_U)$  be the joint transmission policy of the network. Given an initial state of the energy levels  $\mathbf{E}_0 = \mathbf{e}_0 \in$

$\mathcal{E}^U$  and initial scenario  $S_0 = s_0 \in \mathcal{S}$ , we denote the average long-term utility of the data reported by EHS  $u$  to the FC, under the joint transmission policy  $\mu$ , as

$$R_{\mu}^{(u)}(\mathbf{e}_0, s_0) = \liminf_{K \rightarrow \infty} \frac{1}{K} \mathbb{E} \left[ \sum_{k=0}^{K-1} \chi(\gamma_{u,k} > \gamma_{\text{th}}(Y_{u,k})) Q_{u,k} V_{u,k} \times \prod_{i \neq u} (1 - Q_{i,k}) \middle| \mathbf{E}_0 = \mathbf{e}_0, S_0 = s_0, \mu \right], \quad (4)$$

where  $E_{u,k}$  evolves as in (3),  $B_{u,k} \sim \mathcal{B}(\beta(S_k))$ ,  $\{S_k\}$  is a Markov chain, and  $\chi(\cdot)$  is the indicator function. The expectation is taken with respect to  $\{\mathbf{B}_k, S_k, \mathbf{Q}_k, \mathbf{V}_k, \mathbf{Y}_k, \gamma_k\}$ , where  $\mathbf{Y}_k = (Y_{1,k}, Y_{2,k}, \dots, Y_{U,k})$ ,  $\gamma_k = (\gamma_{1,k}, \gamma_{2,k}, \dots, \gamma_{U,k})$ , and, at each instant  $k$ ,  $Q_{u,k} = 1$  with probability  $\mu_u(E_{u,k}, Y_{u,k}, S_k)$  and  $Q_{u,k} = 0$  otherwise. We define the *network utility* as the average long-term aggregate utility of the packets successfully reported to the FC, i.e.,

$$R_{\mu}(\mathbf{e}_0, s_0) = \sum_{u=1}^U R_{\mu}^{(u)}(\mathbf{e}_0, s_0). \quad (5)$$

The goal is to design a joint transmission policy  $\mu$  so as to maximize the network utility, i.e.,

$$\mu^* = \arg \max_{\mu} R_{\mu}(\mathbf{e}_0, s_0). \quad (6)$$

We now establish that  $\mu^*$  has a threshold structure with respect to the observation  $Y_{u,k}$ .

**Proposition 1** *For each energy level  $E_{u,k} \in \mathcal{E}$ , scenario process  $S_k \in \mathcal{S}$ , and for each EHS  $u \in \{1, 2, \dots, U\}$ , there exists a threshold on the utility observation,  $y_{\text{th},u}^*(E_{u,k}, S_k)$ , such that*

$$\mu_u^*(E_{u,k}, Y_{u,k}, S_k) = \chi(Y_{u,k} \geq y_{\text{th},u}^*(E_{u,k}, S_k)). \quad (7)$$

*Proof:* See Appendix A. ■

From Prop. 1, it follows that  $Q_{u,k} = \chi(Y_{u,k} \geq y_{\text{th},u}(E_{u,k}, S_k))$  is optimal, where  $y_{\text{th},u}(E_{u,k}, S_k)$  is the *censoring threshold* on  $Y_{u,k}$ , and is a function of the energy level  $E_{u,k}$  and scenario state  $S_k$ .

We denote the transmission probability of EHS  $u$  in energy level  $e$  and scenario  $s$ , induced by the random observation  $Y_{u,k}$  and by the threshold  $y_{\text{th},u}(e, s)$ , as

$$\eta_u(e, s) = \mathbb{E}[Q_{u,k} | E_{u,k} = e, S_k = s] = \mathbb{P}(Y_{u,k} \geq y_{\text{th},u}(e, s)). \quad (8)$$

We denote the corresponding threshold on the utility observation as  $y_{\text{th}}(\eta_u(e, s)) \triangleq y_{\text{th},u}(e, s)$ . Moreover, we denote the expected utility reported by EHS  $u$  to FC in state  $(e, s)$ , assuming that all the other EHSs remain idle (no collisions occur), as  $g(\eta_u(e, s))$ . This is given by

$$g(\eta_u(e, s)) = \mathbb{E}[\chi(\gamma_{u,k} > \gamma_{\text{th}}(Y_{u,k})) Q_{u,k} V_{u,k} | E_{u,k} = e, S_k = s] \\ = \int_{y_{\text{th},u}(e, s)}^{\infty} \bar{V}(y) f_Y(y) dy. \quad (9)$$

In words,  $g(\eta_u(e, s))$  is the expected reward accrued when only the packets with utility observation larger than  $y_{\text{th},u}(e, s)$  are reported and no collisions occur (however, poor channel fading may still cause the transmission to fail), where  $\eta_u(e, s)$

is the corresponding transmission probability. The function  $g(x)$  has the following properties, which are not explicitly proved here due to space constraints.

**Proposition 2** *The function  $g(x)$  is a continuous, strictly increasing, strictly concave function of  $x$ , with  $g(0) = 0$ ,  $g(1) = \mathbb{E}[\bar{V}(Y_{u,k})]$ ,  $g'(x) \triangleq \frac{dg(x)}{dx} = \bar{V}(y_{\text{th}}(x))$ ,  $\lim_{x \rightarrow 0} g'(x) = \infty$ ,  $\lim_{x \rightarrow 1} g'(x) = 0$ .*

Note that, from Assumption 1, there is a one-to-one mapping between the threshold policy  $\mu_u(e, y, s)$ , the threshold  $y_{\text{th},u}(e, s)$ , and the transmission probability  $\eta_u(e, s)$ . Therefore, without loss of generality, in the following we refer to  $\eta_u$  as the policy of EHS  $u$ . Moreover, we denote the aggregate policy used by all the EHSs in the network as  $\boldsymbol{\eta} = (\eta_1, \eta_2, \dots, \eta_U)$ . The utility for EHS  $u$  under a threshold policy  $\boldsymbol{\eta}$  is then given by

$$R_{\boldsymbol{\eta}}^{(u)}(\mathbf{e}_0, s_0) = \liminf_{K \rightarrow \infty} \frac{1}{K} \mathbb{E} \left[ \sum_{k=0}^{K-1} g(\eta_u(E_{u,k}, S_k)) \times \prod_{i \neq u} (1 - \eta_i(E_{i,k}, S_k)) \middle| \mathbf{E}_0 = \mathbf{e}_0, S_k = s_0, \boldsymbol{\eta} \right], \quad (10)$$

where  $E_{u,k}$  evolves as in (3),  $B_{u,k} \sim \mathcal{B}(\beta(S_k))$  and  $\{S_k\}$  is a Markov chain. The optimization problem under threshold policies can be restated as

$$\boldsymbol{\eta}^* = \arg \max_{\boldsymbol{\eta}} R_{\boldsymbol{\eta}}(\mathbf{e}_0, s_0). \quad (11)$$

Note that the policy  $\eta_u$  implemented by EHS  $u$  yields the following trade-off. By employing a larger transmission probability  $\eta_u(E_{u,k}, S_k)$ :

- EHS  $u$  transmits more often, hence the battery level  $E_{u,k}$  tends to be lower since  $Q_{u,k} \sim \mathcal{B}(\eta_u(E_{u,k}, S_k))$  (see (3)) and battery depletion occurs more frequently for EHS  $u$ , degrading its performance;
- more frequent collisions occur in the channel, hence a lower utility is accrued by all the other EHSs, as apparent from the term  $\prod_{i \neq j} (1 - \eta_i(E_{i,k}, S_k))$  in (10), denoting the probability of successful transmission for EHS  $j$ , which is a decreasing function of  $\eta_u(E_{u,k}, S_k)$ ;
- a larger instantaneous utility is accrued by EHS  $u$ , as apparent from the term  $g(\eta_u(E_{u,k}, S_k))$  in (10).

On the other hand, the opposite effects are obtained by using a smaller transmission probability. The optimal policy  $\boldsymbol{\eta}^*$  thus reflects an optimal trade-off between battery depletion probability, collision probability and instantaneous utility for each EHS.

We consider the following set of *admissible* policies  $\mathcal{U}$ .

**Definition 1** *The set  $\mathcal{U}$  of admissible policies is defined as*

$$\mathcal{U} = \{ \boldsymbol{\eta} : \eta(0, s) = 0, \eta(e_{\max}, s) \in (0, 1], \eta(e, s) \in (0, 1), \\ \forall e = 1, 2, \dots, e_{\max} - 1, \forall s \in \mathcal{S} \}.$$

**Remark 2** *Note that considering only admissible policies  $\boldsymbol{\eta} \in \mathcal{U}$  does not incur a loss of optimality. In fact, under a non-admissible policy, e.g., such that  $\eta(e, s) = 0$  for some  $e > 0$ ,*

then the energy levels below “ $e$ ” are never visited (since the EHS never transmits in state  $E_{u,k} = e$ , hence the energy level cannot decrease). It follows that the system is equivalent to another system with a smaller battery with capacity  $\tilde{e}_{\max} = e_{\max} - e$ . Similarly, if  $\eta(e, s) = 1$  for some  $e < e_{\max}$ , then the energy levels above “ $e$ ” are never visited, hence the system is equivalent to another system with a smaller battery with capacity  $\tilde{e}_{\max} = e$ . In both cases, a non-admissible policy results in under-utilization of the battery storage capacity.

It can be shown that the Markov chain  $\{(\mathbf{E}_k, S_k)\}$  under the aggregate policy  $\eta \in \mathcal{U}$  is irreducible. Hence, there exists a unique steady-state distribution,  $\pi_\eta(e, s), e \in \mathcal{E}^U, s \in \mathcal{S}$ , independent of  $(e_0, s)$  [36]. From (10), we thus obtain

$$R_\eta^{(u)} = \sum_{e \in \mathcal{E}^U, s \in \mathcal{S}} \pi_\eta(e, s) g(\eta_u(e, s)) \prod_{i \neq u} (1 - \eta_i(e_i, s)). \quad (12)$$

Note that, given  $S_k$ , the action  $Q_{u,k}$  is based only on  $(E_{u,k}, Y_{u,k})$  and is independent of  $(E_{i,k}, Y_{i,k}), i \neq u$ . However, the transmission decisions over the network are coupled via the common scenario state  $S_k$ . Herein, we use the following assumption:

**Assumption 2** *The scenario process varies slowly over time, at a time scale much larger than battery dynamics.*

Therefore, conditioned on  $S_k = s$ , a steady-state distribution of the energy levels is approached in each interval during which the scenario stays the same. Then, using an approach similar to [5], since the EH process is i.i.d. across EHSs, the energy level of EHS  $u$  is independent of the energy levels of all the other EHSs, so that we can approximate  $\pi_\eta(e, s)$  as

$$\pi_\eta(e, s) = \pi_S(s) \prod_u \pi_{\eta_u}(e_u | s), \quad (13)$$

where  $\pi_{\eta_u}(e_u | s)$  is the steady state probability of the energy level of EHS  $u$  in scenario  $S_k = s$ , given by the following proposition.

**Proposition 3** *Let  $\eta \in \mathcal{U}$  and  $\xi_\eta(e | s) \triangleq \frac{\beta(s)(1-\eta(e, s))}{(1-\beta(s))\eta(e, s)}$ ; then,  $\pi_\eta(e | s), e \in \mathcal{E}, s \in \mathcal{S}$  is given by*

$$\pi_\eta(0 | s) = \frac{1}{1 + \sum_{j=1}^{e_{\max}} \left[ \prod_{f=1}^{j-1} \xi_\eta(f | s) \right] \frac{\beta(s)}{(1-\beta(s))\eta(j, s)}}, \quad (14)$$

$$\pi_\eta(e | s) = \left[ \prod_{f=1}^{e-1} \xi_\eta(f | s) \right] \frac{\beta(s)\pi_\eta(0 | s)}{(1-\beta(s))\eta(e, s)}, e > 0. \quad (15)$$

*Proof:* We apply the balance equations

$$\pi_\eta(e | s) \beta(s) (1 - \eta(e, s)) = \pi_\eta(e + 1 | s) (1 - \beta(s)) \eta(e + 1, s), \quad e = 0, 1, \dots, e_{\max} - 1, \quad (16)$$

from which (15) is obtained by induction on  $e$ . Finally, (14) is obtained by normalization. ■

Letting

$$G(\eta_u | s) = \sum_{e=1}^{e_{\max}} \pi_{\eta_u}(e | s) g(\eta_u(e, s)), \quad (17)$$

$$P(\eta_u | s) = \sum_{e=1}^{e_{\max}} \pi_{\eta_u}(e | s) \eta_u(e, s),$$

we can rewrite (12) as

$$R_\eta^{(u)} = \sum_{s \in \mathcal{S}} \pi_S(s) G(\eta_u | s) \prod_{i \neq u} (1 - P(\eta_i | s)). \quad (18)$$

Eq. (18) can be interpreted as follows.  $G(\eta_u | s)$  is the average long-term reward of EHS  $u$  in scenario  $s$ , assuming that all the other EHSs remain idle, so that no collisions occur.  $P(\eta_i | s)$  is the average long-term transmission probability of EHS  $i$  in scenario  $s$ , so that  $\prod_{i \neq u} (1 - P(\eta_i | s))$  is the steady-state probability that all EHSs  $i \neq u$  remain idle, i.e., the transmission of EHS  $u$  is successful given that it transmits. From (5), the network utility under the aggregate policy  $\eta = (\eta_1, \eta_2, \dots, \eta_U)$  then becomes

$$R_\eta = \sum_{s \in \mathcal{S}} \pi_S(s) R_\eta[s], \quad (19)$$

where we have defined the network utility in scenario  $s$  as

$$R_\eta[s] \triangleq \sum_{u=1}^U G(\eta_u | s) \prod_{i \neq u} (1 - P(\eta_i | s)). \quad (20)$$

#### A. Symmetric control policies

In order to guarantee fairness among the EHSs, we consider only symmetric control policies, i.e., all the EHSs employ the same policy  $\eta_u = \eta, \forall u$ . From (19) and (20), we thus obtain

$$R_\eta = \sum_{s \in \mathcal{S}} \pi_S(s) U G(\eta | s) (1 - P(\eta | s))^{U-1}. \quad (21)$$

The optimization problem (11) over the class of admissible and symmetric policies is stated as

$$\eta^* = \arg \max_{\eta \in \mathcal{U}} \sum_{s \in \mathcal{S}} \pi_S(s) U G(\eta | s) (1 - P(\eta | s))^{U-1}. \quad (22)$$

Moreover, since each term within the sum depends on the policy used in scenario  $s$  only, and is independent of the policies used in all other scenarios, it follows that we can decouple the optimization as

$$\eta^*(\cdot, s) = \arg \max_{\eta(\cdot, s) \in \mathcal{U}} U G(\eta | s) (1 - P(\eta | s))^{U-1}, \forall s \in \mathcal{S}. \quad (23)$$

Therefore, the optimal threshold policy can be determined independently for each scenario  $s$ , rather than jointly. In the next section, we focus on the solution of the optimization problem (23), for a generic scenario state  $s \in \mathcal{S}$ .

We have the following upper bound to  $R_\eta$ .

**Proposition 4** *The network utility  $R_\eta$  under symmetric threshold policies is upper bounded by*

$$R_\eta < \sum_{s \in \mathcal{S}} \pi_S(s) U g(\min\{x^*, \beta(s)\}) (1 - \min\{x^*, \beta(s)\})^{U-1} \triangleq R^{(up)}, \quad (24)$$

where  $x^* \in (0, 1/U)$  uniquely solves  $\bar{V}(y_{\text{th}}(x^*)) (1 - x^*) = (U-1)g(x^*)$ .

*Proof:* Using the strict concavity of  $g(x)$  (Prop. 2) and Jensen's inequality,  $G(\eta | s)$  is upper bounded by  $G(\eta | s) < g(\sum_{e=1}^{e_{\max}} \pi_\eta(e | s) \eta(e, s)) = g(P(\eta | s))$ . Moreover, the EH

operation in scenario  $s$  with EH rate  $\beta(s)$  induces the constraint on the average transmission probability  $P(\eta|s) \leq \beta(s)$ . Therefore, the network utility (21) can be upper bounded by

$$\begin{aligned} R_\eta &< \sum_{s \in \mathcal{S}} \pi_S(s) U g(P(\eta|s)) (1 - P(\eta|s))^{U-1} \\ &\leq \sum_{s \in \mathcal{S}} \pi_S(s) \max_{x \in [0, \beta(s)]} F(x), \end{aligned} \quad (25)$$

where we have defined  $F(x) \triangleq U g(x) (1 - x)^{U-1}$ . The first derivative of  $F(x)$  with respect to  $x$  is given by

$$F'(x) = U(1 - x)^{U-2} [g'(x)(1 - x) - (U - 1)g(x)]. \quad (26)$$

Therefore  $F(x)$  is an increasing function of  $x$  if and only if  $f(x) \triangleq g'(x)(1 - x) - (U - 1)g(x) > 0$ . Using the strict concavity of  $g(x)$ , we have that  $f(x)$  is a strictly decreasing function of  $x$ , with  $\lim_{x \rightarrow 0} f(x) = \infty$  and  $f(1/U) = -[g(1/U) - \frac{1}{U}g'(1/U)] < 0$ . Therefore, there exists a unique  $x^* \in (0, 1/U)$  such that  $f(x^*) = 0$ ,  $f(x) > 0, \forall x \in (0, x^*)$  and  $f(x) < 0, \forall x \in (x^*, 1/U)$ . Equivalently,  $F(x)$  is a strictly increasing function of  $x \in (0, x^*)$  and strictly decreasing function of  $x \in (x^*, 1/U)$ , so that it is maximized by  $x^*$ . Using Prop. 2 and the definition of  $f(x)$ ,  $x^*$  is the unique solution of  $\bar{V}(y_{\text{th}}(x^*)) (1 - x^*) = (U - 1)g(x^*)$ . Then, by restricting the optimization of  $F(x)$  over  $x \in [0, \beta(s)]$  we obtain  $R_\eta \leq \sum_{s \in \mathcal{S}} \pi_S(s) F(\min\{x^*, \beta(s)\})$ . The proposition is thus proved. ■

In Sec. V, we will present a heuristic policy, which will be shown to achieve the upper bound (25) for asymptotically large battery capacity  $e_{\max} \rightarrow \infty$ .

#### IV. OPTIMIZATION AND ANALYSIS OF THRESHOLD POLICIES

The optimization problem (23) for the case of one EHS ( $U = 1$ ) has been studied in detail in [19] and can be solved using the *policy iteration algorithm* (PIA) [37]. However, in general, when  $U > 1$ , (23) cannot be recast as a convex optimization problem, hence we need to resort to approximate solutions. In particular, in order to determine a local optimum of (23), we use a mathematical artifice based on a game theoretic formulation of the multiaccess problem considered in this paper: we model the optimization problem as a game, where it is assumed that each EHS, say  $u$ , is a player which attempts to maximize the common payoff (20) by optimizing its own policy  $\eta_u$ . Note that the EHSs do not behave strategically, but rather attempt to maximize a *common* network utility with respect to their own policy.

We proceed as follows. We first characterize the general Nash equilibrium (NE) of this game, thus defining the policy profile  $\eta^*$ , over the general space of non symmetric policies  $\eta \in \mathcal{U}^U$ . By definition of NE,  $\eta^*$  is the joint policy such that, if EHS  $u$  deviates from the equilibrium solution by using a different policy  $\eta_u \neq \eta_u^*$ , while all other EHSs  $i \neq u$  use the policy achieving the NE,  $\eta_i^*$ , then a smaller network utility  $R_\eta$  is obtained. This equilibrium condition must hold for all EHSs  $u$ . Equivalently, any *unilateral* deviation from the NE  $\eta^*$  (i.e., one and only one EHS deviates from it) yields a smaller network utility  $R_\eta$ . Then, since our focus is on symmetric

policies, we study the existence of a *symmetric NE* (SNE) for this game, i.e., a NE characterized by the symmetry condition  $\eta_u^* = \eta^*, \forall u$ , so that all EHSs employ the same policy. In Theorem 1, we present structural properties of the SNE. In Theorem 2, we show that the SNE is unique, and we provide Algorithm 2 to compute it. In Theorem 3, we prove that the SNE, and thus the policy returned by Algorithm 2, is a local optimum of the original optimization problem (23). Since in this section we consider the optimization problem (23) for a specific scenario state  $s \in \mathcal{S}$ , for notational convenience we neglect the dependence of  $\eta, \pi_\eta, G(\eta_u|s)$  and  $P(\eta_u|s)$  on  $s$ , i.e., we write  $\eta, \pi_\eta(e), G(\eta_u)$  and  $P(\eta_u)$  instead of  $\eta(\cdot, s), \pi_\eta(e|s), G(\eta_u|s)$  and  $P(\eta_u|s)$ , respectively.

If a NE exists for this game (not necessarily symmetric), defined by the policy profile  $\eta^* = (\eta_1^*, \eta_2^*, \dots, \eta_U^*)$ , then any unilateral deviation from  $\eta^*$  yields a smaller network utility, so that the NE must simultaneously solve,  $\forall u$ ,

$$\begin{aligned} \eta_u^* &= \arg \max_{\eta_u \in \mathcal{U}} R_{\eta_u, \eta_{-u}^*}[s] \\ &= \arg \max_{\eta_u \in \mathcal{U}} \left[ G(\eta_u) \prod_{i \neq u} (1 - P(\eta_i^*)) \right. \\ &\quad \left. + (1 - P(\eta_u)) \sum_{n \neq u} G(\eta_n^*) \prod_{i \neq n, u} (1 - P(\eta_i^*)) \right] \\ &= \arg \max_{\eta_u \in \mathcal{U}} \left[ G(\eta_u) - P(\eta_u) \sum_{n \neq u} \frac{G(\eta_n^*)}{1 - P(\eta_n^*)} \right]. \end{aligned} \quad (27)$$

In the second step, using (20), we have explicated the dependence of the network utility on the policy of EHS  $u$ ,  $\eta_u$  and, in the last step, we have divided the argument of the optimization by the term  $\prod_{i \neq u} (1 - P(\eta_i^*))$ , which is independent of  $\eta_u$ , and hence does not affect the optimization.

By further imposing a symmetric policy  $\eta_u^* = \eta^*, \forall u$ , in (27), we obtain the SNE

$$\eta^* = \arg \max_{\eta \in \mathcal{U}} [G(\eta) - \Lambda(\eta^*) P(\eta)], \quad (28)$$

where we have defined

$$\Lambda(\eta) = (U - 1) \frac{G(\eta)}{1 - P(\eta)}. \quad (29)$$

Note that  $\eta^*$  defined in (28) is *simultaneously* optimal for all the EHSs, i.e., by the definition of Nash equilibrium, any *unilateral* deviation of a single EHS from the SNE  $\eta^*$  yields a smaller network utility  $R_\eta$ .<sup>2</sup> Moreover, although the SNE imposes symmetric policies, it represents an equilibrium over the larger space of non-symmetric policies,  $\eta \in \mathcal{U}^U$ , as dictated by the NE in (27). The interpretation of (28) is as follows.  $G(\eta)$  is the reward obtained by EHS  $u$  when all the other EHSs are always idle. The term  $\Lambda(\eta^*)$  can be interpreted as a Lagrange multiplier associated to a constraint on the transmission probability of EHS  $u$ , so as to limit the collisions caused by its transmissions to the other EHSs in the network. The overall objective function in (28) is thus interpreted as the maximization of the individual reward of each node, with a constraint on the average transmission probability to limit the collisions, which are deleterious to network performance. Interestingly, the Lagrange multiplier (29) increases with the number of EHSs  $U$ , so that, the larger the network size,

<sup>2</sup>Note that these unilateral deviations do not preserve the symmetry of the policy, since, by definition, they assume that one and only one EHS deviates from it.

the more stringent the constraint on the average transmission probability of each EHS, as expected. Moreover, the larger  $P(\eta^*)$  or  $G(\eta^*)$ , i.e., the larger the transmission probability or the reward of the other EHSs, the larger the Lagrange multiplier  $\Lambda(\eta^*)$ , thus imposing a stricter constraint on EHS  $u$  to limit its transmissions and reduce collisions. Note that the optimal Lagrange multiplier  $\Lambda(\eta^*)$  is a function of the SNE (28), hence its value optimally balances the power constraints for the EHSs.

In order to solve (28), we address the more general optimization problem, for  $\lambda \geq 0$ ,

$$\eta^{(\lambda)} = \arg \max_{\eta \in \mathcal{U}} Z_\lambda(\eta), \quad (30)$$

where we have defined  $Z_\lambda(\eta) \triangleq G(\eta) - \lambda P(\eta) = \sum_{e=1}^{e_{\max}} \pi_\eta(e) z_\lambda(\eta(e))$  and  $z_\lambda(x) = g(x) - \lambda x$ . Theorem 1 presents structural properties of  $\eta^{(\lambda)}$ , which are thus inherited by the SNE (28).

### Theorem 1

P1)  $\eta^{(\lambda)}$  is unique, i.e.,  $Z_\lambda(\eta^{(\lambda)}) > Z_\lambda(\eta), \forall \eta \neq \eta^{(\lambda)}$ ;

P2)  $\eta^{(\lambda)}$  is continuous in  $\lambda$ ;

P3)  $\eta^{(\lambda)}(e)$  is a strictly increasing function of  $e$ , and  $\eta^{(\lambda)}(e) \in (\eta_L(\lambda), \eta_H(\lambda))$ ,  $\forall e \neq 0$ , where  $\eta_L(\lambda) \in (0, \min\{x_\lambda^*, \beta(s)\})$ ,  $\eta_U(\lambda) \in [\min\{\beta(s), x_\lambda^*\}, x_\lambda^*]^3$  uniquely solve

$$g(\eta_L(\lambda)) + (1 - \eta_L(\lambda)) \bar{V}(y_{\text{th}}(\eta_L(\lambda))) - \frac{z_\lambda(\min\{x_\lambda^*, \beta(s)\})}{\beta(s)} = \lambda,$$

$$g(\eta_U(\lambda)) - \eta_U(\lambda) \bar{V}(y_{\text{th}}(\eta_U(\lambda))) - z_\lambda(\min\{x_\lambda^*, \beta(s)\}) = 0,$$

and  $x_\lambda^* = \arg \max_{x \in [0,1]} z_\lambda(x)$  is the unique solution in  $[0, 1]$  of  $\bar{V}(y_{\text{th}}(x)) = \lambda$ .<sup>4</sup>

*Proof:* See Appendix B. ■

The policy  $\eta^{(\lambda)}$  can be determined numerically using the following PIA, by exploiting the structural properties of Theorem 1.<sup>5</sup>

### Algorithm 1 (PIA for a given scenario state $s \in \mathcal{S}$ )

1) **Initialization:**  $\eta^{[0]} \in \mathcal{U}$  such that  $\eta^{[0]}(e) \in (\eta_L(\lambda), \eta_H(\lambda))$ ,  $\eta^{[0]}(e+1) > \eta^{[0]}(e)$ ,  $\forall e \neq 0$ ;  $\epsilon_{PIA} \ll 1$ ;  $i = 0$ ;

2) **Policy Evaluation:** compute  $Z_\lambda(\eta^{[i]})$  as in (17), where the steady state distribution  $\pi_{\eta^{[i]}}(e)$  is given by Prop. 3. Let  $D_{\eta^{[i]}}(0) = 0$  and compute recursively, for  $e > 0$ ,

$$D_{\eta^{[i]}}(e) = \frac{Z_\lambda(\eta^{[i]}) - z_\lambda(\eta^{[i]}(e-1))}{\beta(s)(1 - \eta^{[i]}(e-1))} + \frac{(1 - \beta(s))\eta^{[i]}(e-1)}{\beta(s)(1 - \eta^{[i]}(e-1))} D_{\eta^{[i]}}(e-1); \quad (31)$$

3) **Policy Improvement:** determine a new policy as

$$\eta^{[i+1]}(e) = \begin{cases} \eta_H(\lambda), & u(e, \lambda) > \eta_H(\lambda), \\ u(e, \lambda), & u(e, \lambda) \in [\eta_L(\lambda), \eta_H(\lambda)], \\ \eta_L(\lambda), & u(e, \lambda) < \eta_H(\lambda), \end{cases} \quad (32)$$

<sup>3</sup>Note that, if  $\beta(s) \geq x_\lambda^*$ , then  $\eta_U(\lambda) = x_\lambda^*$ .

<sup>4</sup>Note that  $x_\lambda^*$ , as defined here, is not to be confused with  $x^*$ , defined in Prop. 4.

<sup>5</sup>In order to avoid confusion with the superscript  $(\lambda)$  for the Lagrange multiplier in (30), we use the superscript  $[i]$  to denote the iteration index of the algorithm.

where  $u(e, \lambda)$  uniquely solves  $\bar{V}(y_{\text{th}}(u(e, \lambda))) = x^{[i]}(e, \lambda)$ , where

$$x^{[i]}(e, \lambda) = \lambda + \beta(s) D_{\eta^{[i]}}(e+1) + (1 - \beta(s)) D_{\eta^{[i]}}(e), e < e_{\max},$$

$$x^{[i]}(e_{\max}, \lambda) = \lambda + (1 - \beta(s)) D_{\eta^{[i]}}(e_{\max}); \quad (33)$$

4) **Termination test:** If  $|Z_\lambda(\eta^{[i+1]}) - Z_\lambda(\eta^{[i]})| < \epsilon_{PIA}$ , return the optimal policy  $\eta^{(\lambda)} = \eta^{[i+1]}$ ; otherwise, update the counter  $i := i + 1$  and repeat from step 2).

**Remark 3** Algorithm 1 can be computed efficiently. In fact, the steady state distribution  $\pi_{\eta^{[i]}}(e)$  in the policy evaluation step 2) can be computed in closed form from Prop. 3, whose complexity scales as  $O(e_{\max})$ .  $D_{\eta^{[i]}}(e)$  is then computed recursively, so that the resulting complexity of step 2) scales as  $O(e_{\max})$ . The policy improvement step 3) requires to compute, for each energy level  $e$ ,  $u(e, \lambda)$  as the unique solution of  $\bar{V}(y_{\text{th}}(u(e, \lambda))) = x^{[i]}(e, \lambda)$ . This can be done efficiently using the bisection method [38]. If the desired accuracy in the evaluation of  $u(e, \lambda)$  is  $\epsilon_u$ , then approximately  $-\log_2(\epsilon_u)$  bisection steps are sufficient, so that the overall complexity of step 2) scales as  $O(-e_{\max} \log_2(\epsilon_u))$ . Steps 2)-3) are then iterated  $N_{PIA}(\epsilon_{PIA})$  times, depending on the desired accuracy  $\epsilon_{PIA}$ . Typically,  $N_{PIA}(\epsilon_{PIA}) \sim 5 \div 10$ , so that the overall complexity of Algorithm 1 scales as  $O(N_{PIA}(\epsilon_{PIA}) e_{\max} (1 - \log_2(\epsilon_u)))$ . For a further discussion on the complexity and convergence of the PIA, please refer to [37] and references therein.

The following proposition states the optimality of Algorithm 1.

**Proposition 5** Algorithm 1 determines the optimal policy  $\eta^{(\lambda)}$  of (30) when  $\epsilon_{PIA} \rightarrow 0$ .

*Proof:* The optimality of the policy iteration algorithm is proved in [37]. The specific forms of the policy evaluation and improvement steps are proved in Appendix C. ■

By comparing (28) and (30),  $\eta^*$  is optimal for (28) if and only if there exists some  $\lambda^* \geq 0$  such that  $\Lambda(\eta^{(\lambda^*)}) = \lambda^*$ . If such  $\lambda^*$  exists, then  $\eta^* = \eta^{(\lambda^*)}$ .

Note that the SNE may not exist in general. This is the case, for instance, if for every symmetric policy  $\eta$ , with  $\eta_u = \eta_i$ ,  $\forall u, i$ , there exists some unilateral deviation from it which improves the network utility, so that the optimality condition (27) may not hold under any symmetric  $\eta$ . Alternatively, the SNE may not be unique in general, i.e., there may exist multiple symmetric policy profiles satisfying the condition (27). Even though the existence and uniqueness of the SNE, solution of (28), are not granted in general, these properties hold for the model considered in this paper, as stated by the following theorem.

**Theorem 2** There exists a unique  $\eta^* \in \mathcal{U}$  solution of (28), i.e.,  $\exists! \eta^* \in \mathcal{U}$  such that

$$G(\eta^*) - \Lambda(\eta^*)P(\eta^*) > G(\eta) - \Lambda(\eta^*)P(\eta), \quad \forall \eta \neq \eta^*.$$

Moreover, under the optimal SNE  $\eta^*$ ,  $P(\eta^*) \leq \min\{\beta(s), \frac{1}{\bar{V}}\}$ .



*Proof:* See Appendix D. ■

Note that  $P(\eta^*) \leq \frac{1}{U}$ . This is due to the concurrent transmissions of the  $U$  EHSs in the network. In fact, the (average long-term) probability of successful data delivery to the FC by EHS  $u$  is given by  $P(\eta)[1 - (1 - P(\eta))^{U-1}]$ . This is a decreasing function of  $P(\eta)$ , for  $P(\eta) > \frac{1}{U}$ , and therefore any policy  $\eta$  such that  $P(\eta) > \frac{1}{U}$  is sub-optimal, resulting in a larger energy expenditure (hence, less energy availability at the EHSs) and fewer successful transmissions to the FC.

In general, the SNE may yield a sub-optimal solution with respect to the original optimization problem over symmetric policies (23). In fact, although the SNE imposes symmetric policies, it is optimal only with respect to *unilateral* deviations from it (which do not preserve symmetry), not symmetric neighborhoods. In contrast, if the policy employed by each EHS is modified by the same quantity, thus enforcing symmetry, then the resulting network utility may improve, so that the SNE may not be the globally optimal symmetric policy. Therefore, it may occur that a symmetric policy is the SNE but is not a local/global optimum of (23). Similarly, the optimal solution of (23),  $\eta^*$ , may not be a SNE. In fact,  $\eta^*$  is such that any deviation from it over the space of *symmetric policies* yields a smaller network utility. However, for such symmetric policy  $\eta^*$ , there may exist some *unilateral* deviations, occurring over the larger space of non-symmetric policies  $\eta \in \mathcal{U}^U$ , yielding an improved network utility, so that  $\eta^*$  may not be the SNE in this case. Even though the SNE, in general, may not be globally/locally optimal with respect to (23), in the following theorem we show that, in fact, it is a *local* optimum of (23), i.e., any *symmetric* deviation in the neighborhood of the SNE yields a smaller network utility. In the special case where the objective function in (23) is a concave (or quasi-concave [39]) function of  $\eta$ , it follows that the SNE is a global optimum for the original optimization problem (23).

**Theorem 3** *The SNE  $\eta^*$  in (28) is a local optimum for (23).*

*Proof:* See Appendix E. ■

To conclude, we present an algorithm to determine the SNE  $\eta^*$  in (28), hence a local optimum of (23), as proved in Theorem 3. In particular, we employ the bisection method [38] to determine the unique  $\lambda^*$  such that  $h(\lambda^*) = 0$ , where we have defined  $h(\lambda) = \Lambda(\eta^{(\lambda)}) - \lambda$ . Such algorithm uses Algorithm 1 to compute  $\eta^{(\lambda)}$ . The SNE  $\eta^*$  is then determined as  $\eta^* = \eta^{(\lambda^*)}$ .

Given lower and upper bounds  $\lambda_{\min}$  and  $\lambda_{\max}$  of  $\lambda^*$ , the bisection method consists in iteratively evaluating the sign of  $h(\lambda)$  for  $\lambda = (\lambda_{\min} + \lambda_{\max})/2$ , based on which  $\lambda_{\min}$  and  $\lambda_{\max}$  of  $\lambda^*$  can be refined, until the desired accuracy is attained. Its optimality is a consequence of Prop. 7 in Appendix D: if  $h(\lambda) > 0$  (respectively,  $h(\lambda) < 0$ ) for some  $\lambda$ , then necessarily  $\lambda < \lambda^*$  ( $\lambda > \lambda^*$ ). The lower bound is initialized as  $\lambda_{\min} = 0$ . Since  $P(\eta^*) \leq \min\{\beta(s), \frac{1}{U}\}$  (Theorem 2), hence  $G(\eta^*) < g(P(\eta^*)) \leq g(\min\{\beta(s), \frac{1}{U}\})$ , from (29)  $\lambda_{\max}$  is initialized as

$$\lambda^* = \Lambda(\eta^{(\lambda^*)}) < \min \left\{ \frac{U-1}{1-\beta(s)} g(\beta(s)), Ug \left( \frac{1}{U} \right) \right\} = \lambda_{\max}. \quad (34)$$

The bisection algorithm, coupled with Algorithm 1 to determine  $\eta^{(\lambda)}$ , is given as follows.

**Algorithm 2 (Optimal  $\lambda^*$  via bisection algorithm)**

- 1) **Initialization:** accuracy of bisection method  $\epsilon_{BM} > 0$  and of the PIA  $\epsilon_{PIA} > 0$ ,  $\lambda_{\min}^{[0]} = 0$  and  $\lambda_{\max}^{[0]}$  as in (34);  $i = 0$ ;
- 2) **Policy optimization:** Let  $\lambda^{[i]} = (\lambda_{\min}^{[i]} + \lambda_{\max}^{[i]})/2$ ; determine  $\eta^{(\lambda^{[i]})}$  using the PIA;
- 3) **Evaluation:** Evaluate  $h(\lambda^{[i]}) = \Lambda(\eta^{(\lambda^{[i]})}) - \lambda^{[i]}$  under the policy  $\eta^{(\lambda^{[i]})}$ ;
- 4) **Update of lower & upper bounds:**

- If  $h(\lambda^{[i]}) \geq 0$ , update the lower and upper bounds as

$$\begin{cases} \lambda_{\min}^{[i+1]} = \lambda^{[i]}, \\ \lambda_{\max}^{[i+1]} = \min\{\lambda_{\max}^{[i]}, \Lambda(\eta^{(\lambda^{[i]})})\}; \end{cases} \quad (35)$$

- If  $h(\lambda^{[i]}) < 0$ , update the lower and upper bounds as

$$\begin{cases} \lambda_{\min}^{[i+1]} = \max\{\lambda_{\min}^{[i]}, \Lambda(\eta^{(\lambda^{[i]})})\}, \\ \lambda_{\max}^{[i+1]} = \lambda^{[i]}; \end{cases} \quad (36)$$

- 5) **Termination test:** If  $\lambda_{\max}^{[i+1]} - \lambda_{\min}^{[i+1]} < \epsilon_{BM}$ , return the optimal policy  $\eta^* = \eta^{(\lambda^{[i]})}$ ; otherwise, update the counter  $i := i + 1$  and repeat from step 2).

**Remark 4** Note that step 4) updates both  $\lambda_{\min}$  and  $\lambda_{\max}$ . This is because  $h(\lambda)$  is a decreasing function of  $\lambda$  and  $\Lambda(\eta^{(\lambda)})$  is a non-increasing function of  $\lambda$  (Prop. 7 in Appendix D). It follows that, if  $h(\lambda) > 0$ , then  $\lambda < \lambda^* = \Lambda(\eta^{(\lambda^*)}) \leq \Lambda(\eta^{(\lambda)})$ , so that the upper bound can be updated as  $\lambda_{\max}^{[i+1]} = \min\{\lambda_{\max}^{[i]}, \Lambda(\eta^{(\lambda^{[i]})})\}$ . Similarly, if  $h(\lambda) < 0$ , then  $\lambda > \lambda^* = \Lambda(\eta^{(\lambda^*)}) \geq \Lambda(\eta^{(\lambda)})$ , so that the lower bound can be updated as  $\lambda_{\min}^{[i+1]} = \max\{\lambda_{\min}^{[i]}, \Lambda(\eta^{(\lambda^{[i]})})\}$ .

**Remark 5** The maximum number of iterations of Algorithm 2 in order to achieve the desired accuracy  $\epsilon_{BM}$  is  $\log_2(\lambda_{\max}^{[0]}/\epsilon_{BM})$ . By combining it with Algorithm 1, the overall complexity to compute the SNE thus scales as  $O(\log_2(\lambda_{\max}^{[0]}/\epsilon_{BM})N_{PIA}(\epsilon_{PIA})e_{\max}(1 - \log_2(\epsilon_u)))$ .

## V. HEURISTIC POLICY

In this section, we evaluate the following heuristic policy, inspired by the upper bound (24),

$$\begin{aligned} \eta^{(H)}(e, s) &= \arg \max_{x \in [0, \beta(s)]} Ug(x)(1-x)^{U-1} \\ &= \min\{x^*, \beta(s)\}, \quad \forall e > 0, \end{aligned} \quad (37)$$

where  $x^* \in (0, 1/U)$  is defined in Prop. 4. From (17), we obtain

$$G(\eta^{(H)}|s) = (1 - \pi_{\eta^{(H)}}(0|s))g(\min\{x^*, \beta(s)\}), \quad (38)$$

$$P(\eta^{(H)}|s) = (1 - \pi_{\eta^{(H)}}(0|s))\min\{x^*, \beta(s)\}, \quad (39)$$

where  $\pi_{\eta^{(H)}}(0)$  can be derived using Prop. 3 as

$$\pi_{\eta^{(H)}}(0|s) = \begin{cases} \frac{1}{1 + \beta(s) \frac{[\frac{\beta(s)(1-x^*)}{(1-\beta(s))x^*}]^{e_{\max}-1}}{\beta(s)-x^*}}, & x^* < \beta(s), \\ \frac{1-\beta(s)}{e_{\max}+1-\beta(s)}, & x^* \geq \beta(s). \end{cases}$$

Note that, by letting  $e_{\max} \rightarrow \infty$ , we obtain  $\pi_{\eta^{(H)}}(0|s) \rightarrow 0$ , hence  $G(\eta^{(H)}|s) \rightarrow g(\min\{x^*, \beta(s)\})$  and  $P(\eta^{(H)}|s) \rightarrow \min\{x^*, \beta(s)\}$ . Therefore, in the limit of asymptotically large battery capacity we obtain  $\lim_{e_{\max} \rightarrow \infty} R_{\eta^{(H)}} = R^{(up)}$ , i.e., the upper bound  $R^{(up)}$  given by Prop. 4 is asymptotically achieved by the heuristic policy. It follows that the heuristic policy is asymptotically optimal for large battery capacity. Note that the heuristic policy is similar to the policies developed in [40] for general EH networks and infinite battery capacity under the assumption of a centralized controller. Therein, it has been shown that the asymptotically optimal power allocation for the EH network is identical to the optimal power allocation for an equivalent non-EH network whose nodes have infinite energy available and employ the same average transmit power as the corresponding nodes in the EH network. Similarly, the heuristic policy (37) maximizes the upper bound on the network utility in scenario  $s$ ,  $Ug(x)(1-x)^{U-1}$ , achievable for asymptotically large battery capacity  $e_{\max} \rightarrow \infty$  under the policy  $\eta(e, s) = x$ ,  $\forall e$ , and replacing the EH operation of the EHSs with an average energy expenditure constraint equal to the average EH rate, so that  $x \in [0, \beta(s)]$  in (37).

We identify the following operational scenarios.

#### A. Energy-limited scenario $x^* \geq \beta(s)$ (hence $1/U \geq \beta(s)$ )

In this case, the expected amount of energy harvested by the EHSs is small compared to the optimal energy expenditure  $x^*$ , which would optimize the overall utility collected at the FC  $Ug(x^*)(1-x^*)^{U-1}$  (under the ideal case of infinite energy availability); it follows that energy is scarce and the shared multiaccess channel is under-utilized. If this holds for all scenarios  $s \in \mathcal{S}$ , we use the bounds

$$\pi_{\eta^{(H)}}(0|s) = \frac{1 - \beta(s)}{e_{\max} + 1 - \beta(s)} < \frac{1}{e_{\max}}, \quad (40)$$

$$P(\eta^{(H)}|s) < \min\{x^*, \beta(s)\} = \beta(s), \quad (41)$$

to obtain

$$\left(1 - \frac{1}{e_{\max}}\right) R^{(up)} < R_{\eta^{(H)}} < R^{(up)}, \quad (42)$$

so that the upper bound is approached with rate inversely proportional to the battery capacity.

#### B. Network-limited scenario $x^* < \beta(s)$

In this case, the expected amount of energy received by the EHSs is large compared to the optimal energy expenditure  $x^*$ , hence energy is abundant and the shared multiaccess channel is the limiting factor, resulting in under-utilization of the harvested energy. If this holds for all scenarios  $s \in \mathcal{S}$ , letting  $\alpha(s) \triangleq \ln\left(\frac{\beta(s)(1-x^*)}{(1-\beta(s))x^*}\right) > 0$ , we use the bounds

$$\begin{aligned} \pi_{\eta^{(H)}}(0|s) &= \frac{1}{1 + \beta(s) \frac{\left[\frac{\beta(s)(1-x^*)}{(1-\beta(s))x^*}\right]^{e_{\max}} - 1}{\beta(s) - x^*}} = \frac{1}{1 + \beta(s) \frac{e^{\alpha(s)e_{\max}} - 1}{\beta(s) - x^*}} \\ &= \frac{1 - x^*/\beta(s)}{e^{\alpha(s)e_{\max}} - x^*/\beta(s)} < \max_{t \in [0, 1]} \frac{1 - t}{e^{\alpha(s)e_{\max}} - t} = e^{-\alpha(s)e_{\max}}, \\ P(\eta^{(H)}|s) &< \min\{x^*, \beta(s)\} = x^*, \end{aligned} \quad (43)$$

to obtain

$$\left(1 - \sum_s \pi_S(s) e^{-\alpha(s)e_{\max}}\right) R^{(up)} < R_{\eta^{(H)}} < R^{(up)}, \quad (44)$$

so that the upper bound is approached exponentially fast.

In the more general case where  $x^* \geq \beta(s)$  for  $s \in \mathcal{S}_{EL} \subseteq \mathcal{S}$  and  $x^* < \beta(s)$  for  $s \in \mathcal{S}_{NL} \equiv \mathcal{S} \setminus \mathcal{S}_{EL}$ , where  $\mathcal{S}_{EL}$  and  $\mathcal{S}_{NL}$  are the sets of scenario states corresponding to energy-limited or network-limited operation, respectively, we thus obtain

$$\begin{aligned} &\left(1 - \sum_{s \in \mathcal{S}_{EL}} \pi_S(s) \frac{1}{e_{\max}} - \sum_{s \in \mathcal{S}_{NL}} \pi_S(s) e^{-\alpha(s)e_{\max}}\right) R^{(up)} \\ &< R_{\eta^{(H)}} < R^{(up)}, \end{aligned} \quad (45)$$

so that a trade-off between inversely decaying and exponentially decaying gap to the upper bound is achieved.

**Remark 6** Due to the resource constraints of EHSs, an efficient implementation of the proposed scheme can be achieved by solving Algorithm 2 at the FC in a centralized fashion. The SNE  $\eta(\cdot, s)$  for a given scenario  $s$  returned by Algorithm 2 can then be broadcast to the EHSs, which in turn store in a look-up table the corresponding values of the threshold  $y_{th}(e, s)$  as a function of the energy level  $e$  and common scenario  $s$ . The complexity of Algorithm 2 is discussed in Remarks 3 and 5. Alternatively, a lower complexity is obtained by employing the heuristic scheme (37). In this case, the EHSs need to store only a threshold value  $y_{th}^{(H)}(s)$ , independent of the current energy level  $e$ . As shown analytically by (45) and numerically in Sec. VI, such scheme achieves near-optimal performance for sufficiently large battery capacity ( $e_{\max} \simeq 10$ ). Algorithm 2 or the heuristic scheme needs to be computed sporadically at the FC, whenever the environmental or network conditions change (utility distribution  $f_V(\cdot)$  and  $f_{Y|V}(\cdot|v)$ , average EH rate  $\beta(s)$ , battery storage capacity  $e_{\max}$ , e.g., due to battery degradation effects [5], or network size  $U$ ).

## VI. NUMERICAL RESULTS

In this section, we present some numerical results for different settings, by varying the battery capacity  $e_{\max} \in \{1, 10\}$  and the EH rate  $\beta \in \{1/U, 0.1, 0.01\}$ . We consider a static scenario setting  $S_k = s, \forall k$ , and set  $\rho(y) = 0, \forall y$ , so that packet losses are due only to collisions, but not to fading. Note that in this case the choice of the fading model only affects the form of the expected utility  $\bar{V}(y)$  in (2). We model  $V_{u,k}$  as an exponential random variable with unit mean, with pdf  $f_V(v) = e^{-v}, v \geq 0$ , and assume that it is observed perfectly, i.e.,  $Y_{u,k} = V_{u,k}$ . From (8) and (9), we obtain  $g(x) = x(1 - \ln x)$ . We remark that other choices of the distribution of  $V_{u,k}$  and of  $Y_{u,k}$  only affect the shape of the function  $g(x)$ , but do not yield any additional insights with respect to the specific case considered. We evaluate the performance of the following policies: the SNE derived via Algorithm 2; the heuristic policy (HEUR) analyzed in Sec. V; the *energy-balanced policy* (EBP), where each EHS, in each slot, transmits with probability  $\beta$  equal to the average EH rate; and the *network-balanced policy* (NBP), where each EHS, in

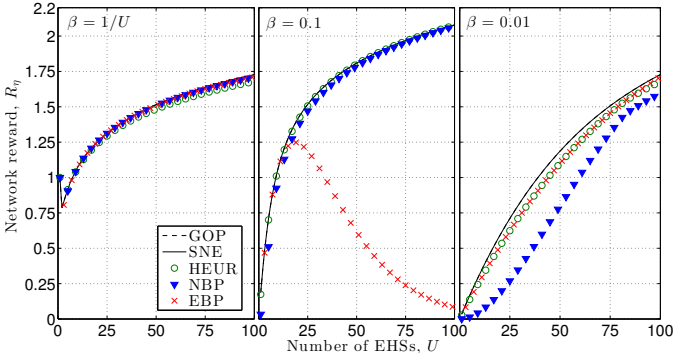


Figure 2. Network utility (10) for different EH rates  $\beta \in \{1/U, 0.1, 0.01\}$ , as a function of the network size  $U$ .  $e_{\max} = 1$ .

each slot, transmits with probability  $1/U$  so as to maximize the expected throughput. For  $e_{\max} = 1$ , we plot also the *globally optimal policy* (GOP) solution of (23). In fact, in this case, the policy is defined by  $\eta(1) \in [0, 1]$ , and can be easily optimized via an exhaustive search, thus serving as an upper bound for the case  $e_{\max} = 1$ . On the other hand, for  $e_{\max} = 10$ , we plot the upper bound (UB) given by Prop. 4. Note that, in general, both SNE and HEUR are sub-optimal with respect to (23), hence GOP and UB provide upper bounds for the cases  $e_{\max} = 1$  and  $e_{\max} = 10$ , respectively, against which the performance of SNE can be evaluated.

In Figs. 2 and 3, we plot the network utility (10), as a function of the number of EHSs  $U$ , for the cases  $e_{\max} = 1$  and  $e_{\max} = 10$ , respectively. We note that, when  $e_{\max} = 1$ , SNE attains the upper bound given by GOP, hence it is optimal for this case. On the other hand, the performance degradation of SNE with respect to UB for the case  $e_{\max} = 10$  is within 3%, hence SNE is near-optimal for this case (note that UB is approached only asymptotically for  $e_{\max} \rightarrow \infty$ ). This suggests that the local optimum achieved by the SNE closely approaches the globally optimal policy. A performance degradation is incurred by HEUR with respect to SNE. In particular, HEUR performs within 18% of GOP for  $e_{\max} = 1$ , and within 9% of UB for  $e_{\max} = 10$ , hence it provides a good heuristic scheme, which achieves both satisfactory performance and low complexity. Note also that HEUR does not require knowledge of the battery energy level, hence it is suitable for those scenarios where the state of the battery is unknown or imperfectly observed [15]. By comparing EBP and NBP, we notice that the former performs better than the latter in the energy-limited scenario  $\beta \leq 1/U$ , and worse in the network-limited scenario  $\beta > 1/U$ .<sup>6</sup> In fact, EBP leads to over-utilization of the channel in the network-limited scenario  $\beta > 1/U$ , hence to frequent collisions and poor performance. On the other hand, NBP matches the transmissions to the channel, without taking into account the energy constraint induced by the EH mechanism, hence it performs poorly in the energy-limited scenario  $\beta \leq 1/U$ , where energy is scarce.

<sup>6</sup>Note that the scenario  $x^* < \beta \leq 1/U$  is indeed *network-limited* (see Sec. V-A). However, both EBP and NBP, unlike HEUR, transmit with probability larger than  $x^*$  in this case, so that their operation is energy-limited (see Sec. V-B). This case is thus included in the *energy-limited scenario* for convenience.

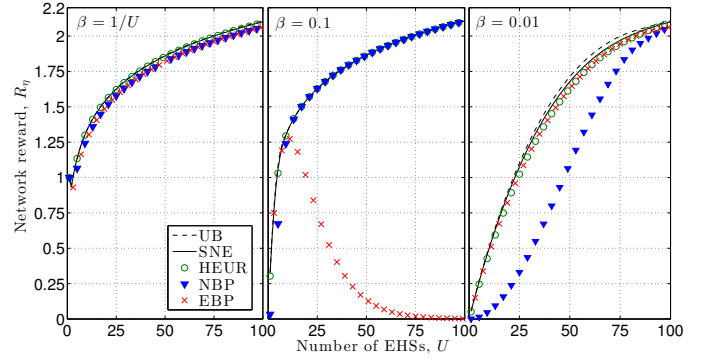


Figure 3. Network utility (10) for different EH rates  $\beta \in \{1/U, 0.1, 0.01\}$ , as a function of the network size  $U$ .  $e_{\max} = 10$ .

Interestingly, in all cases considered, the network utility under the SNE increases with the number of EHSs  $U$ . This behavior is due to the strict concavity of  $g(x)$ , such that a diminishing return is associated to a larger transmission probability  $x$ . Therefore, the smaller the number of EHSs  $U$ , the more the transmission opportunities for each EHS, but the smaller the marginal gain, so that the network utility decreases. For  $U < 10$ , SNE achieves the best performance when  $\beta = 1/U$ , since more energy is available to the EHSs than for  $\beta \in \{0.1, 0.01\}$ . In contrast, for  $U > 10$  and  $e_{\max} = 10$ , SNE performs best in both cases  $\beta = 1/U$  and  $\beta = 0.1$  corresponding to the network-limited scenario, despite a larger energy availability in the latter case. This is due to the fact that, as proved in Theorem 2, under the SNE,  $P(\eta^*) \leq \min\{\beta, 1/U\} = 1/U$ , hence the performance bottleneck is due to the number of EHSs in the network, rather than to the energy availability. In the case  $\beta = 0.1$ , on average,  $\beta - P(\eta^*) \geq \beta - 1/U$  energy quanta are lost in each slot due to overflow, in order to limit the collisions. In contrast, when  $\beta = 0.01$  (energy-limited scenario), we have  $P(\eta^*) \leq \min\{\beta, 1/U\} = \beta$  for all values of  $U$  considered, hence the performance bottleneck is energy availability. A different trend is observed when  $e_{\max} = 1$ . In this case, for  $U > 10$ , SNE performs significantly better in scenario  $\beta = 0.1$  than in scenario  $\beta = 1/U$ . This follows from the fact that, when  $e_{\max} = 1$ , whenever an EHS transmits, its battery becomes empty, hence it enters a recharge phase, with expected duration  $1/\beta = U$ , during which it is inactive. In contrast, the recharge phase in scenario  $\beta = 0.1$  is faster and equals  $10 < U$  slots, so that each EHS becomes more quickly available for data transmission. This phenomenon is negligible when  $e_{\max} = 10$ , since the larger battery capacity minimizes the risk of energy depletion, resulting in the impossibility to transmit for the EHSs.

In Fig. 4, we plot the optimal Lagrange multiplier  $\lambda^* = \Lambda(\eta^*)$  versus the number of EHSs  $U$ . We notice that the larger the number of EHSs, the larger  $\lambda^*$ . In fact, for a given policy  $\eta$ , the larger  $U$ , the more frequent the collisions. A larger  $\lambda^*$  thus balances this phenomenon by penalizing the average transmission probability  $P(\eta^*)$  in (28), and in turn forces each EHS to transmit more sparingly, so as to accommodate the transmissions of more nodes in the network. Moreover, the larger the EH rate  $\beta$ , the larger  $\lambda^*$ . In fact, the larger  $\beta$ ,

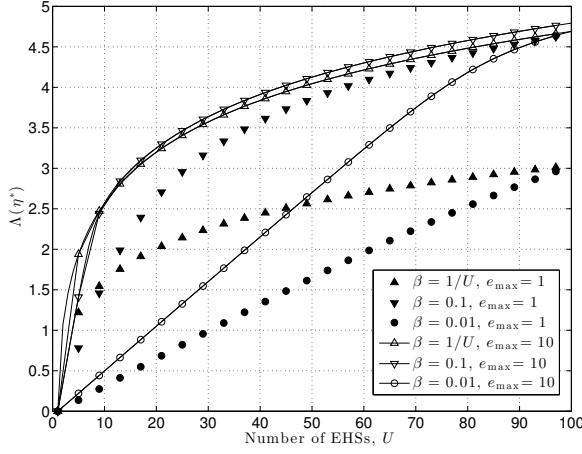


Figure 4.  $\Lambda(\eta^*)$  under the SNE  $\eta^*$  for different values of the battery capacity  $e_{\max} \in \{1, 10\}$  and for different EH rates  $\beta \in \{1/U, 0.1, 0.01\}$ , as a function of the network size  $U$ .

the larger the energy availability for each EHS, which could, in principle, transmit more frequently and, at the same time, cause more collisions. The effect of a larger  $\beta$  (having more transmissions, hence more collisions in the system) is thus balanced by a larger  $\lambda^*$ , which penalizes transmissions.

## VII. CONCLUSIONS

In this paper, we have considered a WSN of EHSs which randomly access a collision channel, to transmit packets of random utility to a common FC, based on local information about the current energy availability and an estimate of the utility of the current data packet. We have studied the problem of designing decentralized random access policies so as to maximize the overall network utility, defined as the average long-term aggregate network utility of the data packets successfully reported to the FC. We have formulated the optimization problem as a game, where each EHS unilaterally maximizes the network utility, and characterized the symmetric Nash equilibrium (SNE). We have proved the existence and uniqueness of the SNE, showing that it is a local optimum of the original optimization problem, and we have derived an algorithm to compute it, and a heuristic policy which is asymptotically optimal for large battery capacity. Finally, we have presented some numerical results, showing that the proposed SNE typically achieves near-optimal performance with respect to the globally optimal policy, at a fraction of the complexity, and we have identified two operational regimes, the *energy-limited* and *network-limited* scenarios, where the bottleneck of the system is given by the energy and the shared multiaccess scheme, respectively.

## APPENDIX A: PROOF OF PROP. 1

Let  $\mu_u$  be a stationary randomized policy used by EHS  $u$ , and  $\mathcal{R}_{\mu_u}$  be the set of stationary randomized policies which induce, in each energy level  $e_u \in \mathcal{E}$  and scenario  $s \in \mathcal{S}$ , the same transmission probability as  $\mu_u$ , defined with respect to

the utility observation, i.e.,  $\forall e_u \in \mathcal{E}, \forall s \in \mathcal{S}, \forall \tilde{\mu}_u \in \mathcal{R}_{\mu_u}$ ,

$$\mathbb{E}[\tilde{\mu}_u(e_u, Y_u, s)] = \mathbb{E}[\mu_u(e_u, Y_u, s)]. \quad (46)$$

Then, since  $\mu_u \in \mathcal{R}_{\mu_u}$ , from (5) the network utility satisfies the inequality

$$R_{\mu_u, \mu_{-u}}(\mathbf{e}_0, s_0) \leq \max_{\tilde{\mu}_u \in \mathcal{R}_{\mu_u}} R_{\tilde{\mu}_u, \mu_{-u}}(\mathbf{e}_0, s_0). \quad (47)$$

We now show that the maximizer of the right hand side of (47) has a threshold structure with respect to the utility observation. From (4) and (5), for any  $\tilde{\mu}_u \in \mathcal{R}_{\mu_u}$ , we have

$$\begin{aligned} R_{\tilde{\mu}_u, \mu_{-u}}(\mathbf{e}_0, s_0) &= \liminf_{K \rightarrow \infty} \frac{1}{K} \sum_{k=0}^{K-1} \sum_{\mathbf{e} \in \mathcal{E}^U, s \in \mathcal{S}} \mathbb{P}(\mathbf{E}_k = \mathbf{e}, S_k = s | \mathbf{e}_0, s_0, \tilde{\mu}_u, \mu_{-u}) \\ &\quad \times \sum_n \mathbb{E} \left[ \chi(\gamma_{n,k} > \gamma_{\text{th}}(Y_{n,k})) Q_{n,k} V_{n,k} \right. \\ &\quad \left. \times \prod_{i \neq n} (1 - Q_{i,k}) \middle| \mathbf{E}_k = \mathbf{e}, S_k = s, \tilde{\mu}_u, \mu_{-u} \right]. \end{aligned}$$

Now, using the fact that  $Q_{n,k} = 1$  with probability  $\mu_n(E_{n,k}, Y_{n,k}, S_k)$ , independent of the energy levels and utility observations of the other EHSs, and  $(V_{n,k}, Y_{n,k}, \gamma_{n,k})$  is independent of  $\mathbf{E}_k$  and  $S_k$ , and i.i.d. across EHSs, we obtain

$$\begin{aligned} &\mathbb{E} \left[ \chi(\gamma_{n,k} > \gamma_{\text{th}}(Y_{n,k})) Q_{n,k} V_{n,k} \right. \\ &\quad \left. \times \prod_{i \neq n} (1 - Q_{i,k}) \middle| \mathbf{E}_k = \mathbf{e}, S_k = s, \tilde{\mu}_u, \mu_{-u} \right] \\ &= \mathbb{E} [\mu(e_n, Y_{n,k}, s) \chi(\gamma_{n,k} > \gamma_{\text{th}}(Y_{n,k})) V_{n,k}] \prod_{i \neq n} (1 - \eta_i(e_i, s)) \\ &= \mathbb{E} [\mu(e_n, Y_{n,k}) \bar{V}(Y_{n,k})] \prod_{i \neq n} (1 - \eta_i(e_i, s)), \quad (48) \end{aligned}$$

where we have defined the transmission probability of EHS  $i$  as  $\eta_i(e_i, s) \triangleq \mathbb{E}[\mu_i(e_i, Y_{i,k}, s)]$ , and we have used (2). Therefore, denoting the steady state probability of  $(\mathbf{E}_k, S_k) = (\mathbf{e}, s)$  as  $\pi_{\tilde{\mu}}(\mathbf{e}, s; \mathbf{e}_0, s_0) = \liminf_{K \rightarrow \infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{P}(\mathbf{E}_k = \mathbf{e}, S_k = s | \mathbf{e}_0, s_0, \tilde{\mu})$ , we obtain

$$\begin{aligned} R_{\tilde{\mu}_u, \mu_{-u}}(\mathbf{e}_0, s_0) &= \sum_{\mathbf{e} \in \mathcal{E}^U, s \in \mathcal{S}} \pi_{\tilde{\mu}_u, \mu_{-u}}(\mathbf{e}, s; \mathbf{e}_0, s_0) \quad (49) \\ &\quad \times \sum_{n=1}^U \mathbb{E} [\mu(e_n, Y_{n,k}, s) \bar{V}(Y_{n,k})] \prod_{i \neq n} (1 - \eta_i(e_i, s)). \end{aligned}$$

Using the fact that  $Y_{u,k}$  is independent of  $\{\mathbf{E}_k, i = 0, \dots, k\}$ ,  $\forall k$ , and i.i.d. across EHSs, it can be proved by induction on  $k$  that the probability  $\mathbb{P}(\mathbf{E}_k = \mathbf{e}, S_k = s | \mathbf{e}_0, s_0, \tilde{\mu}_u, \mu_{-u})$  depends on  $\tilde{\mu}_u$  only through the expected transmission probability  $\eta_u(e_u, s)$ , which is common to all  $\tilde{\mu}_u \in \mathcal{R}_{\mu_u}$ , and therefore the steady state distribution of  $(\mathbf{E}_k, S_k)$  is the same for all

$\tilde{\mu}_u \in \mathcal{R}_{\mu_u}$ , i.e.,  $\pi_{\tilde{\mu}_u, \mu_u}(\mathbf{e}, s; \mathbf{e}_0, s_0) = \pi_{\mu_u, \mu_u}(\mathbf{e}, s; \mathbf{e}_0, s_0)$ . Therefore, from (47) and (49) we obtain

$$\begin{aligned} R_{\mu}(\mathbf{e}_0, s_0) &\leq \max_{\tilde{\mu}_u \in \mathcal{R}_{\mu_u}} R_{\tilde{\mu}_u, \mu_u}(\mathbf{e}_0, s_0) \\ &= \sum_{e_u \in \mathcal{E}, s \in \mathcal{S}} \left[ \mathbb{E} [\mu^*(e_u, Y_{u,k}, s) \bar{V}(Y_{u,k})] \right. \\ &\quad \times \sum_{\mathbf{e}_{-u} \in \mathcal{E}^{U-1}} \pi_{\mu_u, \mu_u}(\mathbf{e}, s; \mathbf{e}_0, s_0) \prod_{i \neq u} (1 - \eta_i(e_i, s)) \\ &\quad + (1 - \eta_u(e_u, s)) \sum_{\mathbf{e}_{-u} \in \mathcal{E}^{U-1}} \pi_{\mu_u, \mu_u}(\mathbf{e}, s; \mathbf{e}_0, s_0) \\ &\quad \left. \times \sum_{n \neq u} \mathbb{E} [\mu_n(e_n, Y_{n,k}, s) V_{n,k}] \prod_{i \neq n, u} (1 - \eta_i(e_i, s)) \right], \quad (50) \end{aligned}$$

where, for each  $e_u \in \mathcal{E}$ ,  $s \in \mathcal{S}$ , we have defined  $\mu_u^*(e_u, \cdot, s)$  as

$$\begin{aligned} \mu_u^*(e_u, \cdot, s) &= \arg \max_{\tilde{\mu}(e_u, \cdot, s): \mathbb{R}^+ \rightarrow [0, 1]} \mathbb{E} [\tilde{\mu}_u(e_u, Y_{u,k}, s) \bar{V}(Y_{u,k})], \\ \text{s.t. } \mathbb{E} [\tilde{\mu}_u(e_u, Y_{u,k}, s)] &= \eta_u(e_u, s). \quad (51) \end{aligned}$$

Since (51) is a convex optimization problem, we can solve it using the Lagrangian method [39], with Lagrangian multiplier  $\bar{V}(y_{\text{th},u}^*(e_u, s))$  for the constraint  $\mathbb{E} [\tilde{\mu}_u(e_u, Y_{u,k}, s)] = \eta_u(e_u, s)$ , i.e.,

$$\mu_u^*(e_u, \cdot, s) = \arg \max_{\tilde{\mu}(e_u, \cdot, s): \mathbb{R}^+ \rightarrow [0, 1]} \mathbb{E} [\tilde{\mu}_u(e_u, Y_{u,k}, s) (\bar{V}(Y_{u,k}) - \bar{V}(y_{\text{th},u}^*(e_u, s)))].$$

The solution to this optimization problem is given by  $\mu_u^*(e_u, y_{u,k}, s) = \chi(\bar{V}(y_{u,k}) \geq \bar{V}(y_{\text{th},u}^*(e_u, s)))$ , and, since  $\bar{V}(y)$  is a strictly increasing and continuous function of  $y$  (see Assumption 1), we obtain the threshold policy (with respect to the utility observation)  $\mu_u^*(e_u, y_{u,k}, s) = \chi(y_{u,k} \geq y_{\text{th},u}^*(e_u, s))$ , where  $y_{\text{th},u}^*(e_u, s)$  is such that  $\mathbb{E} [\mu_u^*(e_u, Y_{u,k}, s)] = \mathbb{P}(Y_{u,k} \geq y_{\text{th},u}^*(e_u, s)) = \eta_u(e_u, s)$ . The threshold structure in (7) is thus proved. ■

## APPENDIX B: PROOF OF THEOREM 1

Due to space constraints, we do not provide an explicit proof of P1 and P2. Due to the concavity of  $g(x)$ ,  $x_\lambda^* = \arg \max_{x \in [0, 1]} z_\lambda(x)$  is the unique solution of  $g'(x) = \lambda$ . Moreover,  $G(\eta) < g(P(\eta))$  from Jensen's inequality [39], hence  $Z_\lambda(\eta) < z_\lambda(P(\eta)) \leq z_\lambda(\min\{x_\lambda^*, \beta\})$ , where we have used the fact that the EH operation enforces the constraint  $P(\eta) \leq \beta$ .

We first prove that  $\eta^{(\lambda)}(e)$  must be a strictly increasing function of the energy level  $e$  (P3.A), by using a technique we term *transmission transfer*.<sup>7</sup> In the second part, we prove that  $\eta^{(\lambda)}(e) \in (\eta_L(\lambda), \eta_H(\lambda))$ ,  $\forall e > 0$  (P3.B). Let  $\eta_0 \in \mathcal{U}$  be a generic transmission policy which violates P3.A, i.e., there exists  $\epsilon \in \{1, \dots, e_{\max} - 1\}$  such that

$$\eta_0(\epsilon - 1) < \eta_0(\epsilon) \geq \eta_0(\epsilon + 1). \quad (52)$$

Note that, since  $\eta_0 \in \mathcal{U}$ , then  $\eta_0(0) = 0$ ,  $\eta_0(1) > 0$ . Then, such  $\epsilon$  can be found as the smallest energy level  $e \geq 1$  such that  $\eta_0(e + 1) \leq \eta_0(e)$ . If such  $\epsilon$  is not found, then necessarily

<sup>7</sup>As we will see, despite the term *transmission transfer*, this technique does not necessarily conserve the amount of transmissions transferred among states.

$\eta_0(0) < \eta_0(1) < \dots < \eta_0(e) < \dots < \eta_0(e_{\max})$ , hence P3.A is not violated. We now define a new transmission policy,  $\eta_\delta$ , parameterized by  $\delta > 0$ , as:

$$\eta_\delta(e) = \begin{cases} \eta_0(e), & e \in \mathcal{E} \setminus \{\epsilon - 1, \epsilon, \epsilon + 1\} \\ \eta_0(\epsilon - 1) + h(\delta), & e = \epsilon - 1 \\ \eta_0(\epsilon) - \delta, & e = \epsilon \\ \eta_0(\epsilon + 1) + r(\delta), & e = \epsilon + 1, \end{cases} \quad (53)$$

where  $h(\delta) > 0$  and  $r(\delta) > 0$  are continuous functions such that  $h(0) = r(0) = 0$ . Intuitively, policy  $\eta_\delta$  is constructed from the original policy  $\eta_0$  by *transferring* some transmissions from energy state  $\epsilon$  to states  $(\epsilon + 1)$  and  $(\epsilon - 1)$ , whereas transmissions in all other states are left unchanged. Therefore, the new policy  $\eta_\delta$  violates less P3.A, by diminishing the gap  $\eta(\epsilon) - \eta(\epsilon + 1)$  by a quantity  $\delta + r(\delta) > 0$ . The functions  $r(\delta) > 0$  and  $h(\delta) \geq 0$  are uniquely defined as

- 1) if  $\epsilon = 1$ , let  $h(\delta) = 0$  and let  $r(\delta)$  be such that  $\pi_{\eta_\delta}(e_{\max}) = \pi_{\eta_0}(e_{\max})$ ,  $\forall \delta < \kappa$ ,
- 2) if  $\epsilon > 1$ , let  $h(\delta)$  and  $r(\delta)$  be such that

$$\begin{cases} \pi_{\eta_\delta}(e_{\max}) = \pi_{\eta_0}(e_{\max}) \\ \pi_{\eta_\delta}(0) = \pi_{\eta_0}(0) \end{cases}, \quad \forall \delta < \kappa, \quad (54)$$

where  $0 < \kappa \ll 1$  is an arbitrarily small constant, which guarantees an admissible policy  $\eta_\delta \in \mathcal{U}$ , i.e.,  $\eta_\delta(\epsilon - 1) \in (0, 1)$ ,  $\eta_\delta(\epsilon) \in (0, 1)$ ,  $\eta_\delta(\epsilon + 1) \in (0, 1)$ .

From Prop. 3, we have that  $\xi_{\eta_\delta}(e) = \xi_{\eta_0}(e)$ ,  $\forall e > \epsilon + 1$ , and therefore, from (15),

$$\begin{aligned} \pi_{\eta_\delta}(e) &= \frac{1}{\prod_{f=e}^{e_{\max}-1} \xi_{\eta_\delta}(f)} \pi_{\eta_\delta}(e_{\max}) \\ &= \frac{1}{\prod_{f=e}^{e_{\max}-1} \xi_{\eta_0}(f)} \pi_{\eta_0}(e_{\max}) = \pi_{\eta_0}(e), \quad e > \epsilon + 1. \quad (55) \end{aligned}$$

Similarly,  $\xi_{\eta_\delta}(e) = \xi_{\eta_0}(e)$ ,  $\forall e < \epsilon - 1$ , and therefore, from (15) and (54),

$$\begin{aligned} \pi_{\eta_\delta}(e) &= \left[ \prod_{f=1}^{e-1} \xi_{\eta_\delta}(f) \right] \frac{\beta \pi_{\eta_\delta}(0)}{(1 - \beta) \eta_\delta(e)} \\ &= \left[ \prod_{f=1}^{e-1} \xi_{\eta_0}(f) \right] \frac{\beta \pi_{\eta_0}(0)}{(1 - \beta) \eta_0(e)} = \pi_{\eta_0}(e), \quad e < \epsilon - 1. \quad (56) \end{aligned}$$

It follows that  $\pi_{\eta_\delta}(e) = \pi_{\eta_0}(e)$  for  $e \in \mathcal{E} \setminus \{\epsilon - 1, \epsilon, \epsilon + 1\}$ , i.e., the transmission transfer preserves the steady state distribution of visiting the low and high energy states. The derivative of  $Z_\lambda(\eta_\delta)$  with respect to  $\delta$ , computed in  $\delta \rightarrow 0^+$ , is then given by

$$\begin{aligned} \left[ \frac{dZ_\lambda(\eta_\delta)}{d\delta} \right]_{\delta \rightarrow 0^+} &= \left[ \frac{d\pi_{\eta_\delta}(\epsilon - 1)}{d\delta} \right]_{\delta \rightarrow 0^+} z_\lambda(\eta_0(\epsilon - 1)) \\ &\quad + \left[ \frac{d\pi_{\eta_\delta}(\epsilon)}{d\delta} \right]_{\delta \rightarrow 0^+} z_\lambda(\eta_0(\epsilon)) + \left[ \frac{d\pi_{\eta_\delta}(\epsilon + 1)}{d\delta} \right]_{\delta \rightarrow 0^+} z_\lambda(\eta_0(\epsilon + 1)) \\ &\quad + \pi_{\eta_0}(\epsilon - 1) z'_\lambda(\eta_0(\epsilon - 1)) h'(0) - \pi_{\eta_0}(\epsilon) z'_\lambda(\eta_0(\epsilon)) \\ &\quad + \pi_{\eta_0}(\epsilon + 1) z'_\lambda(\eta_0(\epsilon + 1)) r'(0), \quad (57) \end{aligned}$$

where  $h'(0) = \left[ \frac{dh(\delta)}{d\delta} \right]_{\delta \rightarrow 0^+}$  and  $r'(0) = \left[ \frac{dr(\delta)}{d\delta} \right]_{\delta \rightarrow 0^+}$ .

We first consider the case  $\epsilon < e_{\max} - 1$ . The case  $\epsilon = e_{\max} - 1$  requires special treatment and will be discussed in the second

part of the proof. For the case  $\epsilon < e_{\max} - 1$ , we compute the derivative terms above. Note that, using (15) and the fact that  $\pi_{\eta_\delta}(\epsilon - 2) = \pi_{\eta_0}(\epsilon - 2)$  and  $\pi_{\eta_\delta}(\epsilon + 2) = \pi_{\eta_0}(\epsilon + 2)$  (if  $\epsilon = 1$ , we define  $\pi_{\eta_\delta}(-1) = \pi_{\eta_0}(-1) = 0$ ), we have

$$\begin{aligned} \frac{\pi_{\eta_0}(\epsilon - 2)}{\pi_{\eta_0}(\epsilon + 2)} &= \frac{\pi_{\eta_\delta}(\epsilon - 2)}{\pi_{\eta_\delta}(\epsilon + 2)} \\ &= \frac{(1-\beta)^4}{\beta^4} \frac{\eta_\delta(\epsilon - 1)}{1-\eta_\delta(\epsilon - 2)} \frac{\eta_\delta(\epsilon)}{1-\eta_\delta(\epsilon - 1)} \frac{\eta_\delta(\epsilon + 1)}{1-\eta_\delta(\epsilon)} \frac{\eta_\delta(\epsilon + 2)}{1-\eta_\delta(\epsilon + 1)}. \end{aligned} \quad (58)$$

By imposing the normalization  $\sum_e \pi_{\eta_\delta}(e) = \sum_e \pi_{\eta_0}(e) = 1$  and using (15), we have that

$$\begin{aligned} \frac{\pi_{\eta_0}(\epsilon - 1) + \pi_{\eta_0}(\epsilon) + \pi_{\eta_0}(\epsilon + 1)}{\pi_{\eta_0}(\epsilon + 2)} &= \frac{\pi_{\eta_\delta}(\epsilon - 1) + \pi_{\eta_\delta}(\epsilon) + \pi_{\eta_\delta}(\epsilon + 1)}{\pi_{\eta_\delta}(\epsilon + 2)} \\ &= \frac{(1-\beta)\eta_\delta(\epsilon + 2)}{\beta(1-\eta_\delta(\epsilon + 1))} + \frac{(1-\beta)\eta_\delta(\epsilon + 1)}{\beta(1-\eta_\delta(\epsilon))} \frac{(1-\beta)\eta_\delta(\epsilon + 2)}{\beta(1-\eta_\delta(\epsilon + 1))} \\ &\quad + \frac{(1-\beta)\eta_\delta(\epsilon)}{\beta(1-\eta_\delta(\epsilon - 1))} \frac{(1-\beta)\eta_\delta(\epsilon + 1)}{\beta(1-\eta_\delta(\epsilon))} \frac{(1-\beta)\eta_\delta(\epsilon + 2)}{\beta(1-\eta_\delta(\epsilon + 1))}. \end{aligned} \quad (59)$$

By computing the derivatives of (58) and (59) with respect to  $\delta \rightarrow 0^+$ , setting them equal to 0 (since they are constant with respect to  $\delta$ ), and solving with respect to  $r'(0)$  with  $h'(0)$ , we obtain

$$r'(0) = \frac{(1-\beta)\eta_0(\epsilon + 1)[1 - \eta_0(\epsilon + 1)]}{[1 - \eta_0(\epsilon)][\beta + \eta_0(\epsilon)(1 - 2\beta)]}, \quad (60)$$

$$h'(0) = \frac{\beta\eta_0(\epsilon - 1)[1 - \eta_0(\epsilon - 1)]}{\eta_0(\epsilon)[\beta + \eta_0(\epsilon)(1 - 2\beta)]}. \quad (61)$$

Note that, if  $\epsilon = 1$ , then (61) yields a feasible  $h'(0) = 0$ , coherent with (53), therefore such solution holds for  $\epsilon = 1$  as well. Using the fact that  $\pi_{\eta_\delta}(\epsilon - 2) = \pi_{\eta_0}(\epsilon - 2)$ , from (15) we obtain

$$\begin{aligned} \pi_{\eta_\delta}(\epsilon - 2) &= \frac{(1-\beta)\eta_\delta(\epsilon - 1)}{\beta(1-\eta_\delta(\epsilon - 2))} \pi_{\eta_\delta}(\epsilon - 1) \\ &= \frac{(1-\beta)\eta_0(\epsilon - 1)}{\beta(1-\eta_0(\epsilon - 2))} \pi_{\eta_0}(\epsilon - 1) = \pi_{\eta_0}(\epsilon - 2), \end{aligned}$$

and therefore, by computing the derivative of the second term above with respect to  $\delta \rightarrow 0^+$ , and setting it equal to zero (since it is constant with respect to  $\delta$ ), we obtain

$$\begin{aligned} \left[ \frac{d\pi_{\eta_\delta}(\epsilon - 1)}{d\delta} \right]_{\delta \rightarrow 0^+} &= -h'(0) \frac{1}{\eta_0(\epsilon - 1)} \pi_{\eta_0}(\epsilon - 1) \\ &= -\frac{1-\beta}{\beta + \eta_0(\epsilon)(1 - 2\beta)} \pi_{\eta_0}(\epsilon). \end{aligned} \quad (62)$$

Similarly, from (15) and the fact that  $\pi_{\eta_\delta}(\epsilon + 2) = \pi_{\eta_0}(\epsilon + 2)$ , we have that

$$\begin{aligned} \pi_{\eta_\delta}(\epsilon + 2) &= \pi_{\eta_\delta}(\epsilon + 1) \frac{\beta(1-\eta_\delta(\epsilon + 1))}{(1-\beta)\eta_\delta(\epsilon + 2)} \\ &= \pi_{\eta_0}(\epsilon + 1) \frac{\beta(1-\eta_0(\epsilon + 1))}{(1-\beta)\eta_0(\epsilon + 2)} = \pi_{\eta_0}(\epsilon + 2), \end{aligned}$$

and therefore, by computing the derivative of the second term above with respect to  $\delta \rightarrow 0^+$ , and setting it equal to zero, we obtain

$$\begin{aligned} \left[ \frac{d\pi_{\eta_\delta}(\epsilon + 1)}{d\delta} \right]_{\delta \rightarrow 0^+} &= r'(0) \frac{1}{1 - \eta_\delta(\epsilon + 1)} \pi_{\eta_0}(\epsilon + 1) \\ &= \frac{\beta}{\beta + \eta_0(\epsilon)(1 - 2\beta)} \pi_{\eta_0}(\epsilon). \end{aligned} \quad (63)$$

Finally, using the fact that  $\pi_{\eta_\delta}(\epsilon - 1) + \pi_{\eta_\delta}(\epsilon) + \pi_{\eta_\delta}(\epsilon + 1) = \pi_{\eta_0}(\epsilon - 1) + \pi_{\eta_0}(\epsilon) + \pi_{\eta_0}(\epsilon + 1)$  (normalization), we obtain

$$\begin{aligned} \left[ \frac{d\pi_{\eta_\delta}(\epsilon)}{d\delta} \right]_{\delta \rightarrow 0^+} &= -\left[ \frac{d\pi_{\eta_\delta}(\epsilon + 1)}{d\delta} \right]_{\delta \rightarrow 0^+} - \left[ \frac{d\pi_{\eta_\delta}(\epsilon - 1)}{d\delta} \right]_{\delta \rightarrow 0^+} \\ &= \frac{1 - 2\beta}{\beta + \eta_0(\epsilon)(1 - 2\beta)} \pi_{\eta_0}(\epsilon). \end{aligned} \quad (64)$$

By substituting these expressions in (57), we obtain

$$\begin{aligned} \frac{\beta + \eta_0(\epsilon)(1 - 2\beta)}{\pi_{\eta_0}(\epsilon)} \left[ \frac{dZ_\lambda(\eta_\delta)}{d\delta} \right]_{\delta \rightarrow 0^+} &= (1 - 2\beta)[g(\eta_0(\epsilon)) - \lambda\eta_0(\epsilon)] \\ &\quad - (1 - \beta)[g(\eta_0(\epsilon - 1)) - \eta_0(\epsilon - 1)g'(\eta_0(\epsilon - 1))] \\ &\quad + \beta\{g(\eta_0(\epsilon + 1)) + g'(\eta_0(\epsilon + 1))[1 - \eta_\delta(\epsilon + 1)] - \lambda\} \\ &\quad - [\beta + \eta_0(\epsilon)(1 - 2\beta)][g'(\eta_0(\epsilon)) - \lambda] \\ &\triangleq S(\eta_0(\epsilon - 1), \eta_0(\epsilon), \eta_0(\epsilon + 1)). \end{aligned} \quad (65)$$

From the concavity of  $g(x)$ ,  $S(\eta_0(\epsilon - 1), \eta_0(\epsilon), \eta_0(\epsilon + 1))$  is a decreasing function of  $\eta_0(\epsilon + 1)$  and  $\eta_0(\epsilon - 1)$ . Then, since  $\eta_0(\epsilon + 1) \leq \eta_0(\epsilon)$  and  $\eta_0(\epsilon + 1) < \eta_0(\epsilon)$  by hypothesis, we have that

$$S(\eta_0(\epsilon - 1), \eta_0(\epsilon), \eta_0(\epsilon + 1)) \geq S(\eta_0(\epsilon), \eta_0(\epsilon), \eta_0(\epsilon)) = 0,$$

hence  $\left[ \frac{d[G(\eta_\delta) - \lambda P(\eta_\delta)]}{d\delta} \right]_{\delta \rightarrow 0^+} > 0$ . Therefore, there exists a small  $\kappa > \delta > 0$  such that  $\eta_\delta \in \mathcal{U}$  and  $Z_\lambda(\eta_\delta) > Z_\lambda(\eta_0)$ , hence  $\eta_0$  is strictly sub-optimal.

Using the same approach for the case  $\epsilon = e_{\max} - 1$ , and imposing the constraints  $\pi_{\eta_\delta}(\epsilon - 2) = \pi_{\eta_0}(\epsilon - 2)$  and  $\pi_{\eta_\delta}(e_{\max}) = \pi_{\eta_0}(e_{\max})$ , it can be proved that

$$\begin{aligned} \frac{\beta + (1-\beta)\eta_0(\epsilon)}{\pi_{\eta_0}(\epsilon)} \left[ \frac{dZ_\lambda(\eta_\delta)}{d\delta} \right]_{\delta \rightarrow 0^+} &= \frac{\beta}{1-\beta}[g'(\eta_0(\epsilon + 1)) - \lambda] \\ &\quad + (1-\beta)[z_\lambda(\eta_0(\epsilon)) - g(\eta_0(\epsilon - 1)) + \eta_0(\epsilon - 1)g'(\eta_0(\epsilon - 1))] \\ &\quad - [g'(\eta_0(\epsilon)) - \lambda][\beta + (1-\beta)\eta_0(\epsilon)] \\ &\triangleq T(\eta_0(\epsilon - 1), \eta_0(\epsilon), \eta_0(\epsilon + 1)). \end{aligned}$$

Then, using the concavity of  $g(x)$ , we have that  $T(\eta_0(\epsilon - 1), \eta_0(\epsilon), \eta_0(\epsilon + 1))$  is a decreasing function of  $\eta_0(\epsilon + 1)$  and  $\eta_0(\epsilon - 1)$ . Then, since  $\eta_0(\epsilon + 1) \leq \eta_0(\epsilon)$  and  $\eta_0(\epsilon - 1) < \eta_0(\epsilon)$  by hypothesis, we obtain

$$\begin{aligned} T(\eta_0(\epsilon - 1), \eta_0(\epsilon), \eta_0(\epsilon + 1)) &> T(\eta_0(\epsilon), \eta_0(\epsilon), \eta_0(\epsilon)) \\ &= \frac{\beta^2}{1-\beta}[g'(\eta_0(\epsilon)) - \lambda]. \end{aligned} \quad (66)$$

Note that, if  $g'(\eta_0(\epsilon)) - \lambda \geq 0$ , then  $\left[ \frac{d[G(\eta_\delta) - \lambda P(\eta_\delta)]}{d\delta} \right]_{\delta \rightarrow 0^+} > 0$ , hence there exists a small  $\kappa > \delta > 0$  such that  $\eta_\delta \in \mathcal{U}$  and  $G(\eta_\delta) - \lambda P(\eta_\delta) > G(\eta_0) - \lambda P(\eta_0)$ , so that  $\eta_0$  is strictly sub-optimal. On the other hand, it can be proved using the transmission transfer technique that a policy  $\eta$  such that  $g'(\eta(e)) - \lambda < 0$  for some  $e$ , i.e.,  $\eta(e) > x_\lambda^*$ , is strictly sub-optimal. An intuitive explanation is that under such policy, the reward in energy level  $e$  satisfies the inequality  $z_\lambda(\eta(e)) < z_\lambda(x_\lambda^*)$ . Therefore, this policy can be improved by a policy  $\tilde{\eta}$  which employs a smaller transmission probability  $\tilde{\eta}(e) = x_\lambda^* < \eta(e)$  in state  $e$  (under a proper transmission transfer to other states). Such policy yields a larger instantaneous reward  $z_\lambda(\tilde{\eta}(e)) > z_\lambda(\eta(e))$  in state  $e$ ,

and a more conservative energy consumption, thus yielding improved long-term performance. The proof of this part is not included due to space constraints.

We now prove P3.B, i.e.,  $\eta^{(\lambda)}(e) \in (\eta_L(\lambda), \eta_H(\lambda))$ ,  $\forall e > 0$ . Using Prop. 3, the derivative of the steady state distribution with respect to  $\eta(1)$  is given by

$$\begin{aligned} \frac{d\pi_\eta(1)}{d\eta(1)} &= \frac{1}{1-\eta(1)} \pi_\eta(1) \left[ 1 - \frac{\pi_\eta(1)}{\beta} \right], \\ \frac{d\pi_\eta(e)}{d\eta(1)} &= -\pi_\eta(e) \pi_\eta(1) \frac{1}{\beta} \frac{1}{1-\eta(1)}, \quad e > 1, \end{aligned} \quad (67)$$

Then, from (17) we obtain

$$\begin{aligned} \frac{dZ_\lambda(\eta)}{d\eta(1)} &= \frac{d\pi_\eta(1)}{d\eta(1)} z_\lambda(\eta(1)) + \pi_\eta(1) [g'(\eta(1)) - \lambda] \\ &+ \sum_{e=2}^{e_{\max}} \frac{d\pi_\eta(e)}{d\eta(1)} z_\lambda(\eta(e)) = \frac{\pi_\eta(1)}{1-\eta(1)} \left( 1 - \frac{\pi_\eta(1)}{\beta} \right) z_\lambda(\eta(1)) \\ &+ \pi_\eta(1) [g'(\eta(1)) - \lambda] - \frac{\pi_\eta(1)}{\beta(1-\eta(1))} \sum_{e=2}^{e_{\max}} \pi_\eta(e) z_\lambda(\eta(e)). \end{aligned}$$

Using the fact that  $\sum_{e=2}^{e_{\max}} \pi_\eta(e) z_\lambda(\eta(e)) = Z_\lambda(\eta) - \pi_\eta(1) z_\lambda(\eta(1))$ , we obtain

$$\begin{aligned} \frac{dZ_\lambda(\eta)}{d\eta(1)} &= \pi_\eta(1) \frac{1}{1-\eta(1)} z_\lambda(\eta(1)) + \pi_\eta(1) [g'(\eta(1)) - \lambda] \\ &- \frac{\pi_\eta(1)}{1-\eta(1)} \frac{Z_\lambda(\eta)}{\beta} \propto g(\eta(1)) + (1-\eta(1))g'(\eta(1)) - \lambda - \frac{Z_\lambda(\eta)}{\beta} \\ &> g(\eta(1)) + (1-\eta(1))g'(\eta(1)) - \lambda - \frac{z_\lambda(\min\{x_\lambda^*, \beta\})}{\beta} \triangleq L(\eta(1)). \end{aligned}$$

Using the concavity of  $g(x)$ , it can be shown that  $L(\eta(1))$  is a decreasing function of  $\eta(1)$ , with  $\lim_{\eta(1) \rightarrow 0^+} L(\eta(1)) = \infty$ . Moreover, using the fact that  $g'(x_\lambda^*) - \lambda = 0$  if  $x_\lambda^* < \beta$ , and  $z_\lambda(\min\{x_\lambda^*, \beta\}) > z_\lambda(0) = 0$  from the definition of  $x_\lambda^*$  and the concavity of  $g(x)$ , we obtain  $L(\min\{x_\lambda^*, \beta\}) = -z_\lambda(\min\{x_\lambda^*, \beta\}) \frac{1-\beta}{\beta} < 0$ . Therefore, there exists a unique  $\eta_L(\lambda) \in (0, \min\{x_\lambda^*, \beta\})$  that solves  $L(\eta_L(\lambda)) = 0$ . Then, for all  $\eta(1) \leq \eta_L(\lambda)$  we have  $L(\eta(1)) \geq 0$ , hence  $\frac{dZ_\lambda(\eta)}{d\eta(1)} > 0$ , which proves that  $\eta(1) \leq \eta_L(\lambda)$  is strictly suboptimal.

Similarly, the derivative of the steady state distribution with respect to  $\eta(e_{\max})$  is given by

$$\frac{d\pi_\eta(e)}{d\eta(e_{\max})} = \frac{\pi_\eta(e) \pi_\eta(e_{\max})}{\eta(e_{\max})}, \quad e < e_{\max}, \quad (68)$$

$$\frac{d\pi_\eta(e_{\max})}{d\eta(e_{\max})} = -\frac{\pi_\eta(e_{\max})}{\eta(e_{\max})} (1 - \pi_\eta(e_{\max})). \quad (69)$$

Then, from (17) we obtain

$$\begin{aligned} \frac{dZ_\lambda(\eta)}{d\eta(e_{\max})} &= \frac{d\pi_\eta(e_{\max})}{d\eta(e_{\max})} z_\lambda(\eta(e_{\max})) + \pi_\eta(e_{\max}) [g'(\eta(e_{\max})) - \lambda] \\ &+ \sum_{e=1}^{e_{\max}-1} \frac{d\pi_\eta(e)}{d\eta(e_{\max})} z_\lambda(\eta(e)) = -\frac{\pi_\eta(e_{\max})}{\eta(e_{\max})} (1 - \pi_\eta(e_{\max})) z_\lambda(\eta(e_{\max})) \\ &+ \pi_\eta(e_{\max}) [g'(\eta(e_{\max})) - \lambda] + \frac{\pi_\eta(e_{\max})}{\eta(e_{\max})} \sum_{e=1}^{e_{\max}-1} \pi_\eta(e) z_\lambda(\eta(e)). \end{aligned}$$

Note that  $\sum_{e=1}^{e_{\max}-1} \pi_\eta(e) z_\lambda(\eta(e)) = Z_\lambda(\eta) - \pi_\eta(e_{\max}) z_\lambda(\eta(e_{\max}))$ . Substituting we obtain

$$\begin{aligned} \frac{dZ_\lambda(\eta)}{d\eta(e_{\max})} &= -\frac{\pi_\eta(e_{\max})}{\eta(e_{\max})} z_\lambda(\eta(e_{\max})) + \pi_\eta(e_{\max}) [g'(\eta(e_{\max})) - \lambda] \\ &+ \frac{\pi_\eta(e_{\max})}{\eta(e_{\max})} Z_\lambda(\eta) \propto -g(\eta(e_{\max})) + \eta(e_{\max}) g'(\eta(e_{\max})) + Z_\lambda(\eta) \\ &< -g(\eta(e_{\max})) + \eta(e_{\max}) g'(\eta(e_{\max})) + z_\lambda(\min\{x_\lambda^*, \beta\}) \\ &\triangleq U(\eta(e_{\max})). \end{aligned} \quad (70)$$

Since  $g(x)$  is concave,  $U(\eta(e_{\max}))$  is a decreasing function of  $\eta(e_{\max})$ . If  $x_\lambda^* \leq \beta$ , using the fact that  $g'(x_\lambda^*) = \lambda$ , we obtain  $U(x_\lambda^*) = 0$ . On the other hand, if  $x_\lambda^* > \beta$ , we have that  $U(x_\lambda^*) = -[g(x_\lambda^*) + g'(x_\lambda^*)(\beta - x_\lambda^*) - g(\beta)] < 0$  and  $U(\beta) = \beta[g'(\beta) - \lambda] > 0$ . Therefore, there exists a unique  $\eta_U(\lambda) \in (\min\{\beta, x_\lambda^*\}, x_\lambda^*)$  that solves  $U(\eta_U(\lambda)) = 0$ . Then, for all  $\eta(e_{\max}) \geq \eta_U(\lambda)$  we have  $U(\eta(e_{\max})) \leq 0$ , hence  $\frac{dZ_\lambda(\eta)}{d\eta(e_{\max})} < 0$ , which proves that  $\eta(e_{\max}) \geq \eta_U(\lambda)$  is strictly suboptimal. Finally, by combining P3.A and P3.B, we obtain  $\eta_L(\lambda) < \eta(1) < \eta(2) < \dots < \eta(e_{\max}) < \eta_U(\lambda)$ . The theorem is thus proved. ■

## APPENDIX C: PROOF OF ALGORITHM 1

For the policy evaluation step, we need to prove that the *relative value function* [37], denoted as  $v_\eta(e)$ , takes the form (31), where we have defined  $D_\eta(e) = v_\eta(e) - v_\eta(e-1)$ ,  $e > 0$ ,  $D_\eta(0) = 0$ . By the definition of relative value function [37],  $v_\eta(e)$  is obtained as the unique solution of the linear system

$$v_\eta(0) = 0, \quad v_\eta(e) - \sum_{f \in \mathcal{E}} \mathbb{P}_\eta(f|e) v_\eta(f) = z_\lambda(\eta(e)) - Z_\lambda(\eta), \quad e \in \mathcal{E},$$

where  $\mathbb{P}_\eta(f|e)$  is the transition probability from energy level  $e$  to energy level  $f$ , under policy  $\eta$ . In particular, by explicating  $\mathbb{P}_\eta(f|e)$ , the linear system above for  $e < e_{\max}$  is equivalent to

$$D_\eta(e+1) = \frac{Z_\lambda(\eta) - z_\lambda(\eta(e))}{\beta(1-\eta(e))} + \frac{(1-\beta)\eta(e)}{\beta(1-\eta(e))} D_\eta(e), \quad (71)$$

proving the recursion (31). In the policy improvement step, we solve the optimization problem

$$\eta^{[i+1]}(e) = \arg \max_{\eta(e) \in (\eta_L, \eta_H)} z_\lambda(\eta(e)) + \sum_f \mathbb{P}_\eta(f|e) v^{[i]}(f), \quad \forall e \neq 0.$$

In particular, by explicating the transition probability  $\mathbb{P}_\eta(f|e)$ , we obtain

$$\eta^{[i+1]}(e) = \arg \max_{\eta(e) \in (\eta_L, \eta_H)} g(\eta(e)) - x^{[i]}(e, \lambda) \eta(e), \quad (72)$$

where  $x^{[i]}(e, \lambda)$  is given by (33). The solution (32) directly follows by exploiting the strict concavity of  $g(x)$  and the fact that  $g'(x) = \bar{V}(y_{\text{th}}(x))$ . ■

## APPENDIX D: PROOF OF THEOREM 2

In order to prove the theorem, we need the following propositions.

**Proposition 6**  $P(\eta^{(\lambda)})$  is a non-increasing function of  $\lambda$ , and  $P(\eta^{(0)}) > 0$ ,  $\lim_{\lambda \rightarrow \infty} P(\eta^{(\lambda)}) = 0$ .

*Proof:* Assume by contradiction that  $\lambda_1 > \lambda_2$  and  $P(\eta^{(\lambda_2)}) < P(\eta^{(\lambda_1)})$ . Then we have

$$Z_{\lambda_2}(\eta^{(\lambda_2)}) \geq Z_{\lambda_1}(\eta^{(\lambda_1)}) + (\lambda_1 - \lambda_2)P(\eta^{(\lambda_1)}) \quad (73)$$

$$> Z_{\lambda_1}(\eta^{(\lambda_1)}) + (\lambda_1 - \lambda_2)P(\eta^{(\lambda_2)}) \quad (74)$$

$$\geq G(\eta^{(\lambda_2)}) - \lambda_1 P(\eta^{(\lambda_2)}) + (\lambda_1 - \lambda_2)P(\eta^{(\lambda_2)}) \quad (75)$$

$$= G(\eta^{(\lambda_2)}) - \lambda_2 P(\eta^{(\lambda_2)}),$$

where in (73)-(74) we have used the fact that  $Z_{\lambda_2}(\eta^{(\lambda_2)}) \geq Z_{\lambda_2}(\eta^{(\lambda_1)})$  and then the hypothesis, in (75) the fact that  $Z_{\lambda_1}(\eta^{(\lambda_1)}) \geq Z_{\lambda_1}(\eta^{(\lambda_2)})$ . By comparing (74) and (75), we obtain the contradiction  $Z_{\lambda_2}(\eta^{(\lambda_2)}) > Z_{\lambda_2}(\eta^{(\lambda_2)})$ . For the second part of the proposition, since  $\eta^{(0)} \neq \{\text{idle policy}\}$ , then  $P(\eta^{(0)}) > 0$ . Since  $P(\eta^{(\lambda)}) < \eta_U(\lambda) \leq x_\lambda^*$  (P3 of Theorem 1), where  $x_\lambda^* = \arg \max_x [g(x) - \lambda x]$ , in the limit  $\lambda \rightarrow \infty$  we obtain  $x_\lambda^* \rightarrow 0$ , hence  $P(\eta^{(\lambda)}) \rightarrow 0$ . The proposition is thus proved. ■

**Proposition 7**  $\Lambda(\eta^{(\lambda)})$  is a continuous, non-increasing function of  $\lambda$ , for  $\lambda \geq 0$ , with limits  $\Lambda(\eta^{(0)}) \in (0, \infty)$  and  $\lim_{\lambda \rightarrow \infty} \Lambda(\eta^{(\lambda)}) = 0$ .  $h(\lambda) \triangleq \Lambda(\eta^{(\lambda)}) - \lambda$  is a continuous, strictly decreasing function of  $\lambda$ , with limits  $h(0) > 0$  and  $\lim_{\lambda \rightarrow \infty} h(\lambda) = -\infty$ .

*Proof:* The continuity of  $\Lambda(\eta^{(\lambda)})$  follows from the continuity of  $\eta^{(\lambda)}$  (P2 of Theorem 1) and the definition of  $\Lambda(\eta)$  in (29). For  $\lambda = 0$  we have  $\eta^{(0)} = \arg \max_{\eta \in \mathcal{U}} G(\eta)$ . Then, we obtain

$$\Lambda(\eta^{(0)}) = (U - 1) \frac{G(\eta^{(0)})}{1 - P(\eta^{(0)})} \in (0, \infty). \quad (76)$$

$\Lambda(\eta^{(0)})$  is positive and bounded since  $0 < G(\eta^{(0)}) \leq g(1) < \infty$  and  $P(\eta^{(0)}) \leq \beta < 1$ . On the other hand, for  $\lambda \rightarrow \infty$ , we have  $\eta^{(\lambda)} \rightarrow 0$ , hence  $G(\eta^{(\lambda)}) \rightarrow 0$ ,  $P(\eta^{(\lambda)}) \rightarrow 0$  and  $\lim_{\lambda \rightarrow \infty} \Lambda(\eta^{(\lambda)}) = 0$ . To conclude, we prove that  $\Lambda(\eta^{(\lambda)})$  is a non-increasing function of  $\lambda$ , i.e.,  $\Lambda(\eta^{(\lambda_1)}) \leq \Lambda(\eta^{(\lambda_2)})$  for  $\lambda_1 > \lambda_2 \geq 0$ . Using (29), this is true if and only if

$$G(\eta^{(\lambda_2)})(1 - P(\eta^{(\lambda_1)})) - G(\eta^{(\lambda_1)})(1 - P(\eta^{(\lambda_2)})) \geq 0.$$

Equivalently, by rearranging the terms,

$$A(\lambda_1, \lambda_2) \triangleq Z_{\lambda_2}(\eta^{(\lambda_2)})[1 - P(\eta^{(\lambda_1)})] - Z_{\lambda_1}(\eta^{(\lambda_1)})[1 - P(\eta^{(\lambda_2)})] + \lambda_2 P(\eta^{(\lambda_2)})[1 - P(\eta^{(\lambda_1)})] - \lambda_1 P(\eta^{(\lambda_1)})[1 - P(\eta^{(\lambda_2)})] \geq 0.$$

Using the fact that  $\eta^{(\lambda_2)}$  is optimal for (30) for  $\lambda = \lambda_2$ , hence  $Z_{\lambda_2}(\eta^{(\lambda_2)}) \geq Z_{\lambda_2}(\eta^{(\lambda_1)})$ , a sufficient condition which guarantees that  $A(\lambda_1, \lambda_2) \geq 0$  is that

$$A(\lambda_1, \lambda_2) \geq Z_{\lambda_2}(\eta^{(\lambda_1)})[1 - P(\eta^{(\lambda_1)})] - Z_{\lambda_1}(\eta^{(\lambda_1)})[1 - P(\eta^{(\lambda_2)})] + \lambda_2 P(\eta^{(\lambda_2)})(1 - P(\eta^{(\lambda_1)})) - \lambda_1 P(\eta^{(\lambda_1)})(1 - P(\eta^{(\lambda_2)})) \geq 0.$$

By rearranging the terms, it can be readily verified that the second term above is equivalent to

$$[P(\eta^{(\lambda_2)}) - P(\eta^{(\lambda_1)})][G(\eta^{(\lambda_1)}) + \lambda_2(1 - P(\eta^{(\lambda_1)}))] \geq 0,$$

which holds from Prop. 6. The second part of the proposition readily follows. ■

Using Prop. 7, we have that there exists a unique  $\lambda^* \in (0, \infty)$  such that  $h(\lambda^*) = 0$ , i.e.,  $\Lambda(\eta^{(\lambda^*)}) = \lambda^*$ . Under such

$\lambda^*$ ,  $\eta^{(\lambda^*)}$  is optimal for (28), and  $\eta^{(\lambda^*)}$  is unique from P1 of Theorem 1.

We now prove that under the SNE,  $P(\eta^{(\lambda^*)}) \leq \frac{1}{U}$  ( $P(\eta^{(\lambda^*)}) \leq \beta$  as shown in the proof of Prop. 4). This is true if  $P(\eta^{(0)}) \leq \frac{1}{U}$ , since  $P(\eta^{(\lambda)})$  is a non-increasing function of  $\lambda$  (Prop. 6). Now, assume that  $P(\eta^{(0)}) > \frac{1}{U}$ . Since  $\lim_{\lambda \rightarrow \infty} P(\eta^{(\lambda)}) = 0$  (Prop. 6), there exists  $\hat{\lambda} \in (0, \infty)$  such that  $P(\eta^{(\hat{\lambda})}) = \frac{1}{U}$ . For such  $\hat{\lambda}$ , from (30) we have

$$G(\eta^{(\hat{\lambda})}) - \hat{\lambda} \frac{1}{U} = Z_{\hat{\lambda}}(\eta^{(\hat{\lambda})}) = \max_{\eta \in \mathcal{U}} Z_{\hat{\lambda}}(\eta) \geq Z_{\hat{\lambda}}(\text{idle}) = 0, \quad (77)$$

where “idle” denotes the policy where the EHS remains always idle, resulting in  $G(\text{idle}) = P(\text{idle}) = 0$ . The inequality follows from the fact that  $Z_{\hat{\lambda}}(\eta)$  is evaluated under the specific policy “idle.” Then, using (77), (29) and the fact that  $P(\eta^{(\hat{\lambda})}) = \frac{1}{U}$  by definition of  $\hat{\lambda}$ , we obtain

$$UG(\eta^{(\hat{\lambda})}) - \hat{\lambda} = (U - 1) \frac{G(\eta^{(\hat{\lambda})})}{1 - 1/U} - \hat{\lambda} = \Lambda(\eta^{(\hat{\lambda})}) - \hat{\lambda} \geq 0.$$

Therefore, we obtain  $\Lambda(\eta^{(\hat{\lambda})}) \geq \hat{\lambda}$ . Since  $\Lambda(\eta^{(\lambda)}) - \lambda$  is a decreasing function of  $\lambda$  (Prop. 7) and the Lagrange multiplier  $\lambda^*$  of the SNE must satisfy  $\Lambda(\eta^{(\lambda^*)}) - \lambda^* = 0$ , necessarily  $\hat{\lambda} \leq \lambda^*$ . Finally, using Prop. 6, we obtain  $P(\eta^{(\hat{\lambda})}) = \frac{1}{U} \geq P(\eta^{(\lambda^*)})$ , thus proving the theorem. ■

#### APPENDIX E: PROOF OF THEOREM 3

From Theorem 1, we have that  $\eta^*(e) \in (\eta_L(\lambda^*), \eta_H(\lambda^*))$  where  $0 < \eta_L(\lambda^*) < \eta_H(\lambda^*) < 1$ . It follows that  $\eta^* \in \text{int}(\mathcal{U})$ . Therefore, since  $\eta^*$  is globally optimal for the optimization problem (28) and it is unique, the gradient with respect to  $\eta$  of the objective function in (28), computed in  $\eta^*$  and denoted as  $\Delta_{\eta^*}(\cdot)$ , is equal to zero, and its Hessian with respect to  $\eta$ , computed in  $\eta^*$  and denoted as  $\mathbf{H}_{\eta^*}(\cdot)$ , is negative definite, i.e., for the SNE  $\eta^*$  we have

$$\Delta_{\eta^*}(G(\eta)) - \Lambda(\eta^*)\Delta_{\eta^*}(P(\eta)) = \mathbf{0}, \quad (78)$$

$$\mathbf{H}_{\eta^*}(G(\eta)) - \Lambda(\eta^*)\mathbf{H}_{\eta^*}(P(\eta)) \prec 0. \quad (79)$$

On the other hand, the gradient of the network utility (21) is given by

$$\Delta_{\eta}(R_{\eta}) = U(1 - P(\eta))^{U-1} \Delta_{\eta}(G(\eta)) - U(U - 1)G(\eta)(1 - P(\eta))^{U-2} \Delta_{\eta}(P(\eta)). \quad (80)$$

The Hessian matrix of (21) is then obtained by further computing the gradient of (80), yielding

$$\begin{aligned} \mathbf{H}_{\eta}(R_{\eta}) &= U(1 - P(\eta))^{U-1} \mathbf{H}_{\eta}(G(\eta)) \\ &\quad - U(U - 1)(1 - P(\eta))^{U-2} \Delta_{\eta}(G(\eta)) \Delta_{\eta}(P(\eta))^T \\ &\quad - U(U - 1)(1 - P(\eta))^{U-2} \Delta_{\eta}(P(\eta)) \Delta_{\eta}(G(\eta))^T \\ &\quad - U(U - 1)G(\eta)(1 - P(\eta))^{U-2} \mathbf{H}_{\eta}(P(\eta)) \\ &\quad + U(U - 1)(U - 2)G(\eta)(1 - P(\eta))^{U-3} \Delta_{\eta}(P(\eta)) \Delta_{\eta}(P(\eta))^T. \end{aligned} \quad (81)$$

By computing (80) under the SNE  $\eta^*$ , using (29) and substituting (78) in (80), we obtain  $\Delta_{\eta^*}(R_{\eta}) = \mathbf{0}$ . Moreover,



since  $\mathbf{H}_{\eta^*}(G(\eta)) \prec \Lambda(\eta^*)\mathbf{H}_{\eta^*}(P(\eta))$  from (79), from (81) we obtain

$$\begin{aligned} \mathbf{H}_{\eta^*}(R_{\eta}) &\prec -U^2(U-1)G(\eta^*)(1-P(\eta))^{U-3} \\ &\times \Delta_{\eta^*}(P(\eta))\Delta_{\eta^*}(P(\eta))^T \preceq 0. \end{aligned}$$

Therefore,  $\mathbf{H}_{\eta^*}(R_{\eta}) \prec 0$  and  $\Delta_{\eta^*}(R_{\eta}) = \mathbf{0}$ , hence  $\eta^*$  is a local optimum for (23). ■

## REFERENCES

- [1] N. Michelusi and M. Zorzi, "Optimal random multiaccess in energy harvesting Wireless Sensor Networks," in *IEEE International Conference on Communications Workshops (ICC)*, June 2013, pp. 463–468.
- [2] M. Miozzo, D. Zordan, P. Dini, and M. Rossi, "SolarStat: Modeling photovoltaic sources through stochastic Markov processes," in *IEEE International Energy Conference (ENERGYCON)*, May 2014, pp. 688–695.
- [3] D. Gunduz, K. Stamatiou, N. Michelusi, and M. Zorzi, "Designing intelligent energy harvesting communication systems," *IEEE Communications Magazine*, vol. 52, no. 1, pp. 210–216, Jan. 2014.
- [4] J. A. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Pervasive Computing*, vol. 4, pp. 18–27, Jan. 2005.
- [5] N. Michelusi, L. Badia, R. Carli, L. Corradini, and M. Zorzi, "Energy Management Policies for Harvesting-Based Wireless Sensor Devices with Battery Degradation," *IEEE Transactions on Communications*, vol. 61, no. 12, pp. 4934–4947, Dec. 2013.
- [6] C. Renner and V. Turau, "CapLibrate: self-calibration of an energy harvesting power supply with supercapacitors," in *Proc. Int. conf. on Architecture of Comput. Syst. (ARCS)*, Hannover, Germany, Feb. 2010, pp. 1–10.
- [7] M. Gorlatova, A. Wallwater, and G. Zussman, "Networking Low-Power Energy Harvesting Devices: Measurements and Algorithms," *IEEE Transactions on Mobile Computing*, vol. 12, no. 9, pp. 1853–1865, Sept 2013.
- [8] D. Anthony, P. Bennett, M. C. Vuran, M. B. Dwyer, S. Elbaum, A. Lacy, M. Engels, and W. Wehtje, "Sensing through the continent: towards monitoring migratory birds using cellular sensor networks," in *Proceedings of the 11th international conference on Information Processing in Sensor Networks (ISPN)*, vol. 12, Apr. 2012, pp. 329–340.
- [9] K. Lin, J. Yu, J. Hsu, S. Zahedi, D. Lee, J. Friedman, A. Kansal, V. Raghunathan, and M. Srivastava, "Helimote: enabling long-lived sensor networks through solar energy harvesting," in *Proc. ACM SenSys*, 2005.
- [10] K. Kar, A. Krishnamurthy, and N. Jaggi, "Dynamic node activation in networks of rechargeable sensors," *IEEE/ACM Transactions on Networking*, vol. 14, pp. 15–26, Feb. 2006.
- [11] D. Niyato, E. Hossain, and A. Fallahi, "Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: performance analysis and optimization," *IEEE Transactions on Mobile Computing*, vol. 6, pp. 221–236, Feb. 2007.
- [12] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Transactions on Embedded Computing Systems*, vol. 6, no. 4, Sep. 2007.
- [13] M. Gatzianas, L. Georgiadis, and L. Tassioulas, "Control of wireless networks with rechargeable batteries," *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, pp. 581–593, Feb. 2010.
- [14] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1326–1336, Apr. 2010.
- [15] N. Michelusi, L. Badia, and M. Zorzi, "Optimal Transmission Policies for Energy Harvesting Devices With Limited State-of-Charge Knowledge," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 3969–3982, Nov 2014.
- [16] N. Michelusi, K. Stamatiou, and M. Zorzi, "Transmission policies for energy harvesting sensors with time-correlated energy supply," *IEEE Transactions on Communications*, vol. 61, no. 7, pp. 2988–3001, July 2013.
- [17] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with Energy Harvesting Nodes in Fading Wireless Channels: Optimal Policies," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1732–1743, Aug. 2011.
- [18] D. Del Testa, N. Michelusi, and M. Zorzi, "On Optimal Transmission Policies for Energy Harvesting Devices: the case of two users," in *Proc. of the 10th Intern. Symposium on Wireless Communication Systems (ISWCS)*, Aug 2013, pp. 1–5.
- [19] N. Michelusi, K. Stamatiou, and M. Zorzi, "On optimal transmission policies for energy harvesting devices," in *Information Theory and Applications Workshop (ITA)*, 2012, Feb. 2012, pp. 249–254.
- [20] J. Lei, R. Yates, and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replenishable sensors," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 547–551, Feb. 2009.
- [21] R. A. Valles, A. G. Marques, and J. G. Sueiro, "Optimal selective forwarding for energy saving in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 164–175, Jan. 2011.
- [22] N. Jaggi, K. Kar, and A. Krishnamurthy, "Rechargeable sensor activation under temporally correlated events," *Springer Wireless Networks (WINET)*, vol. 15, pp. 619–635, Jul. 2009.
- [23] A. Seyed and B. Sikdar, "Energy Efficient Transmission Strategies for Body Sensor Networks with Energy Harvesting," *IEEE Transactions on Communications*, vol. 58, no. 7, pp. 2116–2126, July 2010.
- [24] F. Iannello, O. Simeone, and U. Spagnolini, "Medium Access Control Protocols for Wireless Sensor Networks with Energy Harvesting," *IEEE Transactions on Communications*, vol. 60, no. 5, pp. 1381–1389, May 2012.
- [25] F. Iannello, O. Simeone, P. Popovski, and U. Spagnolini, "Energy group-based dynamic framed ALOHA for wireless networks with energy harvesting," in *46th Annual Conference on Information Sciences and Systems (CISS)*, March 2012.
- [26] V. Sharma, U. Mukherji, and V. Joseph, "Efficient energy management policies for networks with energy harvesting sensor nodes," in *46th Annual Allerton Conference on Communication, Control, and Computing*, Sept 2008, pp. 375–383.
- [27] Z. A. Eu and H.-P. Tan, "Probabilistic polling for multi-hop energy harvesting wireless sensor networks," in *IEEE International Conference on Communications (ICC)*, June 2012, pp. 271–275.
- [28] X. Fafoutis and N. Dragoni, "Adaptive media access control for energy harvesting Wireless sensor networks," in *Ninth International Conference on Networked Sensing Systems (INSS)*, June 2012, pp. 1–4.
- [29] L. Lin, N. B. Shroff, and R. Srikant, "Energy-Aware Routing in Sensor Networks: A Large System Approach," *Ad-Hoc Networks*, vol. 5, no. 6, pp. 818–831, May 2007.
- [30] S. Chen, P. Sinha, N. Shroff, and C. Joo, "A Simple Asymptotically Optimal Joint Energy Allocation and Routing Scheme in Rechargeable Sensor Networks," *IEEE/ACM Transactions on Networking*, vol. 22, no. 4, pp. 1325–1336, Aug 2014.
- [31] C. Tapparello, O. Simeone, and M. Rossi, "Dynamic Compression-Transmission for Energy-Harvesting Multihop Networks With Correlated Sources," *IEEE/ACM Transactions on Networking*, vol. 22, no. 6, pp. 1729–1741, Dec 2014.
- [32] L. Huang and M. Neely, "Utility Optimal Scheduling in Energy-Harvesting Networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 4, pp. 1117–1130, Aug 2013.
- [33] N. Michelusi and U. Mitra, "Cross-Layer Design of Distributed Sensing-Estimation With Quality Feedback – Part I: Optimal Schemes," *IEEE Transactions on Signal Processing*, vol. 63, no. 5, pp. 1228–1243, March 2015.
- [34] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Inc., New York, 2006.
- [35] C. K. Ho, P. D. Khoa, and P. C. Ming, "Markovian models for harvested energy in wireless communications," in *Proc. IEEE Int. Conf. on Communication Systems (ICCS)*, 2010, pp. 311–315.
- [36] D. J. White, *Markov Decision Processes*. Wiley, 1993.
- [37] D. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific, Belmont, Massachusetts, 2005.
- [38] R. L. Burden and J. D. Faires, *Numerical Analysis, 9th Edition*. Cengage Learning, 2011.
- [39] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [40] N. Zlatanov, R. Schober, and Z. Hadzi-Velkov, "Asymptotically Optimal Power Allocation for Energy Harvesting Communication Networks," *CoRR*, vol. abs/1308.2833, 2013. [Online]. Available: <http://arxiv.org/abs/1308.2833>



**Nicolò Michelusi** (S'09, M'13) received the B.Sc. (with honors), M.Sc. (with honors) and Ph.D. degrees from the University of Padova, Italy, in 2006, 2009 and 2013, respectively, and the M.Sc. degree in Telecommunications Engineering from the Technical University of Denmark in 2009, as part of the T.I.M.E. double degree program. In 2011, he was at the University of Southern California, Los Angeles, USA, and, in Fall 2012, at Aalborg University, Denmark, as a visiting research scholar. He is currently a post-doctoral research fellow at the Ming Hsieh

Department of Electrical Engineering, University of Southern California, USA. His research interests lie in the areas of wireless networks, stochastic optimization, distributed estimation and modeling of bacterial networks. Dr. Michelusi serves as a reviewer for the IEEE Transactions on Communications, IEEE Transactions on Wireless Communications, IEEE Transactions on Information Theory, IEEE Transactions on Signal Processing, IEEE Journal on Selected Areas in Communications, and IEEE/ACM Transactions on Networking.



**Michele Zorzi** (S'89, M'95, SM'98, F'07) received his Laurea and PhD degrees in electrical engineering from the University of Padova in 1990 and 1994, respectively. During academic year 1992-1993 he was on leave at UCSD, working on multiple access in mobile radio networks. In 1993 he joined the faculty of the Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy. After spending three years with the Center for Wireless Communications at UCSD, in 1998 he joined the School of Engineering of the University of Ferrara,

Italy, where he became a professor in 2000. Since November 2003 he has been on the faculty of the Information Engineering Department at the University of Padova. His present research interests include performance evaluation in mobile communications systems, random access in wireless networks, ad hoc and sensor networks, Internet-of-Things, energy constrained communications and protocols, cognitive networks, and underwater communications and networking.

He was Editor-In-Chief of IEEE Wireless Communications from 2003 to 2005 and Editor-In-Chief of the IEEE Transactions on Communications from 2008 to 2011, and is currently the Editor-in-Chief of the new IEEE Transactions on Cognitive Communications and Networking. He has been an Editor for several journals and a member of the Organizing or the Technical Program Committee for many international conferences. He was also guest editor for special issues in IEEE Personal Communications ("Energy Management in Personal Communications Systems"), IEEE Network ("Video over Mobile Networks") and IEEE Journal on Selected Areas in Communications ("Multimedia Network Radios," "Underwater Wireless Communications and Networking" and "Energy Harvesting and Energy Transfer"). He served as a Member-at-Large of the Board of Governors of the IEEE Communications Society from 2009 to 2011, and is currently its Director of Education.